

PROF. A. K. SAHA
NUCLEAR PHYSICS DIVISION
SAHA INSTITUTE OF NUCLEAR PHYSICS
82, Acharya Pratulla Chandra Basu,
CALCUTTA-9.

B. B. Ray.
Calcutta

DR. AJIT KUMAR SAHA.
125, SOUTHERN AVENUE
CALCUTTA-29

WAVE-MECHANICS

WAVE-MECHANICS

BY

ARNOLD SOMMERFELD

PROFESSOR OF THEORETICAL PHYSICS AT THE UNIVERSITY OF MUNICH
FOREIGN MEMBER OF THE ROYAL SOCIETY

TRANSLATED FROM THE GERMAN EDITION BY

HENRY L. BROSE, M.A.

D. PHIL., F. INST. P.

SUPPLEMENTARY VOLUME TO
ATOMIC STRUCTURE AND SPECTRAL LINES

WITH THIRTY-THREE FIGURES



METHUEN & CO. LTD.
6 ESSEX STREET, W.C.
LONDON

Originally Published in German under the title
"Atombau und Spektrallinien : Wellenmechanischer Ergänzungsband"

This Translation First Published in 1930

PRINTED IN GREAT BRITAIN

THE AUTHOR'S PREFACE TO THE GERMAN EDITION

THE author believes that he is fulfilling a general wish in attempting in the present supplementary volume to depict the recent developments which since 1924 to 1926 (thanks to the work of L. de Broglie, Heisenberg, Schrödinger) have transformed the external aspect of atomic physics. The fact that the inner content of the theory, that is, the quantitative assertions that can be tested by experiment, has for the most part survived this process of regeneration is manifest to those acquainted with the subject. The new development does not signify a radical change but a welcome evolution of the existing theory, while many fundamental points are classified and made more precise.

I have called this volume a "wave-mechanical" supplement, because for practical manipulation Schrödinger's methods are obviously superior to the specifically "quantum-mechanical" methods. On the other hand, however, I have left no doubt that the general ideas that have led Heisenberg to enunciate quantum-mechanics are also indispensable for the elaboration of wave-mechanics. The original standpoint of Schrödinger, that transitions are to occur only between co-existing states, is clearly too narrow and does not accurately fit the facts. I have therefore taken over into wave-mechanics the equal treatment of states and transitions,—as is done by Heisenberg from the very beginning,—in particular in deriving the frequency-condition and the rules of polarisation and intensity in § 5, Chap. I. This, of course, denotes that I am renouncing the more definite wave-kinematic objective, set up by Schrödinger and de Broglie, and am sacrificing pictorial representation to formalism. In wave-mechanics the electron still remains a point-charge ultimately, and the light-quantum a point-like centre of energy. But the dualism between the light-quantum and the light-wave extends into the corpuscular

region; beside the electron-corpuscule we have the electron-wave with all the accumulative experimental evidence to confirm it.

During several semesters I have endeavoured in my university lectures to make clear to my class and to myself the chief results of wave-mechanics in as simple a form as possible. We found that in all cases that permitted of complete integration the "polynomial method" is adequate and leads to the final analytical form. The "method of generating functions," although often very elegant, is here replaced by directly applying the conditions of orthogonality. My purpose in the present volume was to bring out such points of view clearly. A more significant simplification is that which we have been able to achieve in Dirac's theory of the electron.

The form in which this supplementary volume is presented follows directly on the 4th German edition of *Atombau und Spektrallinien*. Chapter I contains the fundamental conceptions and their application to elementary problems, whereas Chapter II deals with the more difficult calculations of perturbations and diffraction and—the most difficult problem of all—the relativistic theory of the electron. The present volume can of course also be regarded as a supplement to the English edition, *Atomic Structure and Spectral Lines*, to which references are made where possible.

As in the previous editions of the main volume I have kept in view both the theoretical and the experimental physicist as readers. I have therefore restricted myself to such problems as may claim direct physical interest. The general speculations of the Theory of Transformations of Probabilities receive far too small a share of space, as do the fundamental questions of Indeterminacy and Observability. Concerning these more general subjects, I understand that other accounts by authoritative writers are about to appear. My wish was to preserve the original character of my book, and I therefore kept my attention as much as possible, on concrete questions. I would gladly have treated afresh the systematic arrangement of spectra from the point of view of wave-mechanics. But there was neither sufficient space nor time for this; moreover, the theory of the relativistic electron will have to be further investigated before these things can become sufficiently clear.

In §§ 4 and 7 of Chapter II on the photo-electric and Compton effects I have had the pleasure of the collaboration of my colleague, Mr. F. Kirchner. I am indebted to Dr. S. Bochner for many

mathematical references and suggestions. But my special thanks are due to my faithful associate, Dr. K. Bechert, who has given me his untiring help not only in all the technical questions of the printing and in all the details of the manuscript, but has also worked out many points of the text independently, particularly in the final paragraph on electron spin, so that if I had not had his devoted help it would have scarcely been possible to bring the book to completion at the appointed time.

A. SOMMERFELD

MUNICH
August, 1928

Prof.- A. K. SAHA.
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS.
92, Acharya Pratulla Chandra Road,
CALCUTTA-9.

Prof:- A. K. SAHA.
NUCLEAR PHYSICS DIVISION.
SAHA INSTITUTE OF NUCLEAR PHYSICS.
92, Acharya Prafulla Chandra Road,
CALCUTTA-9.

TRANSLATOR'S PREFACE

SINCE the appearance of the last edition of the author's *Atombau und Spektrallinien*, which is generally recognised as the standard work on the subject, many striking advances have been made owing to the birth and growth of quantum- and wave-mechanics. What is the attitude of the author towards these new theories? Professor Sommerfeld has endeavoured in the present supplementary volume to answer this question and to bring his treatment of the subject completely up to date. The success of his attempt is clearly demonstrated by the highly appreciative way in which the German original has been received. The great number of original researches which have contributed to these recent advances often appear little connected outwardly; by-paths which are ventured upon and later abandoned tend to confuse the average reader whose mathematical attainments may be insufficient to allow him to form a just estimate of such contributions. The masterly and brilliant way in which Professor Sommerfeld has sifted the highly abstract material and brought into prominence what is of practical importance to the physicist will be recognised by all who wish to follow the thread of progress of atomic physics under his inspiring guidance. He has described in a particularly convenient form the mathematical machinery that is necessary for an understanding of the theory.

The present English edition has had the advantage of being accurately checked and revised by the author. Many additions and alterations have been made; for example, the last section of Chapter I has a different form from that which it was given in the German edition; several oversights in the German edition have also been corrected. A particularly welcome feature is that Professor Sommerfeld has corrected and supplemented the section on the Photo-electric Effect, on which much interest is focussed at the present time, by adding three Notes, rich in content, which

form the conclusion of the volume. The English edition may therefore claim to be a stage in advance of the German original.

References to the English edition of the main work are briefly denoted by I placed before the page or chapter quoted. In those cases in which reference is made to passages in the 4th German edition which do not occur in the English version (which was translated from the 3rd German edition), the abbreviation *Atombau* precedes the page or chapter quoted.

In the hope of simplifying the reading of original papers on the subject the translator has appended a list of the German expressions which occur in the text, together with the English equivalents used by him. I am indebted to Professor H. T. H. Piaggio of University College, Nottingham, for some helpful suggestions in this connection. The arduous task of proof-reading was very kindly undertaken by Mr. H. F. Biggs, M.A., of the Electrical Laboratory, Oxford, and Mr. E. H. Saayman, M.A., of New College, Oxford, who devoted much time and care to correcting and improving the English rendering. I also wish to express my gratitude to Professor Sommerfeld himself for his friendly interest in the course of the work and for finally checking the proofs.

HENRY L. BROSE

UNIVERSITY COLLEGE, NOTTINGHAM
December, 1929

CONTENTS

	PAGE
THE AUTHOR'S PREFACE TO THE GERMAN EDITION	v
TRANSLATOR'S PREFACE	ix

CHAPTER I

INTRODUCTION TO WAVE-MECHANICS, RUDIMENTS AND SIMPLEST APPLICATIONS . 1-148	
§ 1. Schrödinger's Wave-Equation and the de Broglie Wave-Length	1
§ 2. Explanation of the Mathematical Method. Spherical Harmonics. Bessel Functions	7
§ 3. Oscillator and Rotator. Their Proper Values according to Wave-Mechanics	14
§ 4. The Oscillator according to Quantum-Mechanics	29
§ 5. Completion of Wave-Mechanics. Frequency Condition. Questions of Intensity	38
§ 6. Intensities in the case of Oscillators and Rotators	48
§ 7. The Kepler Problem	59
§ 8. Statistical View of the Continuous Charge-cloud. Complete and Incomplete Shells. Magnetic Moments. Spatial Quantising. Physical Applications (Sphere of Action in Collisions, Supposed Double Refraction in a Magnetic Field, Lattice Forces)	88
§ 9. Relativistic Generalisations. The Principle of Variation. Application to the Kepler Problem	100
§ 10. The Zeeman Effect	119
§ 11. The Molecule as a Symmetrical Top	123
§ 12. Concerning the Relationship between Quantum Conditions in Wave-Mechanics and in the Older Quantum Theory	184

CHAPTER II

PERTURBATIONS, DIFFRACTION PROBLEMS. THE SPINNING ELECTRON . . 144-289	
§ 1. Schrödinger's Theory of Perturbations	144
§ 2. The Stark Effect	153
§ 3. Theory of Dispersion	164
§ 4. Photo-electric Effect	177

	PAGE
§ 5. Diffraction at a Single Obstacle. Collision Problems of Two Particles .	192
§ 6. Diffraction of Matter by Gratings. Interference of Electron-Waves in Crystals	205
§ 7. The Compton Effect	212
§ 8. The Helium Problem	225
§ 9. Wave-Mechanical Transcription of Classical Quantities	240
§ 10. Concerning the Nature of the Electron	257
NOTES	290-298
1. On the Continuous Proper Functions in the Hydrogen Spectrum, their Asymptotic Representation and their Normalisation	290
2. Concerning the One-sidedness of Photo-electric Emission. Calculation of the Integrals K_1 and K_2	295
3. The Integration over the Continuous Spectrum Transition from the Proper Function ψ to the Partial Functions ψ^1, ψ^2	296
INDEX OF GERMAN EXPRESSIONS AND THEIR ENGLISH EQUIVALENTS . . .	299
INDEX	300

CHAPTER I

INTRODUCTION TO WAVE-MECHANICS, RUDIMENTS AND SIMPLEST APPLICATIONS

§ 1. Schrödinger's Wave-Equation and the de Broglie Wave-Length

THE antithesis between macroscopic and microscopic events has often been emphasised. For example, the state of a configuration in heat equilibrium looks quite different when regarded microscopically than from the point of view of the kinetic theory of gases. Mechanics and electrodynamics are also macroscopic in origin. To apply them unchanged to the conditions in the atom is to make unjustifiable demands of Nature. Nevertheless important partial successes favoured the extrapolation of these theories to microscopic conditions. The study of the electronic orbits within the atom, which has become so immensely fruitful for our knowledge of the atom, in particular for deciphering spectra, was founded on classical mechanics. And the considerations of the Correspondence Principle, which we required in order to answer questions of intensity and polarisation were derived from classical electrodynamics. To these two classical principles there were added two quantum axioms. We indicate them briefly (see Vol. I,* Chap. IV, § 1, and Chap. I, § 6) by means of the two equations

$$\oint p dq = nh \quad (\text{Quantum condition}) \quad . \quad . \quad . \quad (1)$$

$$h\nu = E_1 - E_2 \quad (\text{Frequency condition}) \quad . \quad . \quad (2)$$

The first defines the favoured or stationary states of the atom (more generally, of the system in question) and distinguishes them by means of an integer n (quantum number). The second determines the radiation emitted, in a transition from one state to another, by means of the corresponding energies E_1 and E_2 .

But there were various indications that the mechanical laws, even when amplified in this way by the quantum theory, did not yield the whole truth. A particularly simple case in which they failed was that of the rotation-bands of molecules. These had to be numbered not by integers but by half-integers (cf. I, Chap. VII, § 2, particularly p. 418) if the facts

* In future, references to Volume I, English edition, will be denoted merely by I.

† In the sequel we prefer to use E in place of the former W in order to indicate that in general we now normalise the energy differently from before, namely, from a rational zero point. Whereas this different normalisation is of no consequence for the present, it will become essential in § 5, and particularly in the generalisation founded on relativity in § 8.

of experiment were to be represented without arbitrary restrictions. Another much discussed example was given by the anomalous Zeeman effect. Landé's g -formula, which can be confirmed with extreme accuracy, requires $n(n+1)$ and $j(j+1)$, etc., in all places where we should expect n^2 and j^2 , etc., according to the classical theory.* For a time one spoke of a "non-mechanical constraint" or of "an ambiguity of the model that defied mechanical description." Further, the hitherto accepted theory failed in questions involving mutual actions, for example, in the case of the helium problem, and indeed also in other cases (higher spectral series), where the fault could not be ascribed to the elusive helium model (cf. I, Chap. III, § 6). It is true that recent investigations have shown that the blame for some of these alleged contradictions to the principles of macro-mechanics is not to be imputed to mechanics but to the structure of the electron. This is so, in particular, in the case of the magneto-mechanical anomaly, which occurs in the Zeeman effect,† of the Paschen-Back effect for the hydrogen atom and of the half-integral values of j in the general structure of the system of series spectra. But, apart from these, there are still other examples of disagreement which point to the antithesis between micro- and macro-mechanics.

How can we arrive at a theory of micro-mechanics appropriate to atomic phenomena without undue arbitrariness? We shall follow Erwin Schrödinger‡ by starting from the comprehensive analytical system of Hamiltonian mechanics. Hamilton developed it in connection with his researches on the geometrical optics of astronomical instruments.§ His ideas were inspired by the undulatory theory of optics, which was coming to the fore at that time (1828-1837).

The wave-theory of optics describes optical phenomena by means of *linear partial differential equations of the second order* and derives the wave-surfaces (surfaces of constant phase) from them. From the point of view of the wave-theory, at least for isotropic media, light-rays may be defined as the orthogonal trajectories of wave-surfaces.

Geometrical or ray-optics, on the other hand, was originally a mechanics of Newtonian light-corpuscles. The light-rays denoted the paths of these particles. Hamilton took over the idea of wave-surfaces from wave-optics and constructed them as the surfaces orthogonal to the rays. If their equation is $S = \text{const.}$, then S satisfies a *partial differential equation of the first order and second degree*. This is Hamilton's partial differential equation of mechanics. S denotes the Hamiltonian

* Cf. Chap. VIII, p. 629, of 4th German edition of *Atombau und Spektrallinien*, or Birtwistle, *New Quantum Mechanics*, Camb. Univ. Press.

† *Atombau* (4th edn.), pp. 620 and 635.

‡ Schrödinger's collected papers on wave-mechanics have been published in English in one volume by Messrs. Blackie & Son, Ltd.; this volume gives the complete German references. The chief papers appeared in the *Annalen der Physik*, Vols. 79, 80, and 81.

§ Besides the references given in Note 7, p. 804, *Atombau*, see also F. Klein, *Entwicklung der Mathematik im 19. Jahrhundert*, Springer 1926, Vol. I, Chap. V.

§ 1. Schrödinger's Wave-Equation, the de Broglie Wave-Length 3

characteristic function, or "action function": this corresponds with our earlier nomenclature in I, page 556. In this way we arrive at the differential equation and the action function for the individual point-mass (the individual light-corpuscle); this method is then easily extended to mechanical systems in general.

We shall now follow Hamilton's course in the reverse direction. Whereas Hamilton, starting from wave-optics, passed by way of ray-optics to the general formulation of macro-mechanics, we shall follow Schrödinger and proceed from macro-mechanics by way of ray- and wave-optics to micro-mechanics. Just as wave-optics is a refinement of ray-optics for dimensions of the order of the wave-length, so we expect to arrive at a micro-mechanics which refines macro-mechanics and may be applied to atomic dimensions.

We begin with the macro-mechanics of the individual point-mass in rectangular co-ordinates. Starting from the energy-equation

$$\frac{m}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) = E - V. \quad (3)$$

(E = energy-constant, V = potential energy, expressed as a function of x, y, z alone). We obtain, according to the general rules of I, page 558, the Hamiltonian differential equation

$$\Delta_1 S = 2m(E - V), \text{ where } \Delta_1 S \equiv \left(\frac{\partial S}{\partial x}\right)^2 + \left(\frac{\partial S}{\partial y}\right)^2 + \left(\frac{\partial S}{\partial z}\right)^2. \quad (4)$$

On the other hand, we write down the differential equation of wave-optics in the form

$$\Delta u = \frac{1}{a^2} \frac{\partial^2 u}{\partial t^2} \text{ where } \Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}. \quad (5)$$

and where u is a rectangular component of the optical field or of the corresponding vector potential, and a the phase-velocity of the light, which in general varies from point to point. We shall, however, at once eliminate the dependence on time, which we do not discuss till § 5, by making the substitution, for monochromatic light,

$$u = \psi e^{i\omega t} \quad (5a)$$

We set

$$\frac{\omega}{a} = k$$

and call k the "wave-number." This term is justified in view of the fact that for a plane wave (cf. the end of this section) k becomes equal to $2\pi/\lambda$, where λ = the wave-length, that is, the spatial periodicity of the plane wave. Thus, in the case of a plane wave, k signifies, in more exact language, the number of wave-lengths that are included in 2π units of length. Further, we introduce the refractive index n against vacuum

4 Chapter I Introduction to Wave-Mechanics

(the indices 0 refer to "vacuum," thus a_0 is equal to the ordinary velocity of light, c):

$$n = \frac{a_0}{a} = \frac{c}{a}, \quad k = nk_0.$$

With these terms defined in this way, it follows from (5) and (5a) that

$$\Delta\psi + n^2 k_0^2 \psi = 0 \quad (6)$$

We accomplish the transition to ray-optics by following an idea of Debye's* as follows. In ray-optics we regard the wave-length λ_0 as small, that is, k_0 as "large" ("small" means infinitely small compared with all the dimensions of the optical apparatus that are involved). We write

$$\psi = A e^{ik_0 S} \quad (6a)$$

and treat A and S as "slowly" varying quantities; that is, in the derivatives

$$\begin{aligned} \frac{\partial \psi}{\partial x} &= \left(ik_0 A \frac{\partial S}{\partial x} + \frac{\partial A}{\partial x} \right) e^{ik_0 S} \\ \frac{\partial^2 \psi}{\partial x^2} &= \left[-k_0^2 A \left(\frac{\partial S}{\partial x} \right)^2 + ik_0 A \frac{\partial^2 S}{\partial x^2} + 2ik_0 \frac{\partial A}{\partial x} \frac{\partial S}{\partial x} + \frac{\partial^2 A}{\partial x^2} \right] e^{ik_0 S} \end{aligned}$$

we neglect all the lower powers of k_0 in comparison with the highest occurring power. If we then cancel k_0^2 , (6) becomes

$$\Delta_1 S = n^2 \quad (7)$$

This is the differential equation of the "eikonal," the characteristic function for ray-optics. Comparison with (4) (a dimensional difficulty that occurs will be discussed presently) gives the formal relation

$$n^2 = 2m(E - V) \quad (8)$$

If we translate Hamiltonian mechanics into the language of ray-optics, we see that it operates with a refractive index that varies according to the measure of V : this gives us something similar to the curvilinear course of rays through the layers of the earth's atmosphere. We insert this value of n in the differential equation (6) of wave-optics, and thus obtain in addition to the picture of mechanics based on ray-optics one based on wave-optics and represented by

$$\Delta\psi + 2m(E - V)k_0^2 \psi = 0 \quad (9)$$

A remark on dimensions must here be added. In equation (4) S has the dimensions of an action (erg sec), whereas in (7) it has the dimensions of a length. Hence the comparison of (7) with (4) leads to a formula with wrong dimensions (on the left is a pure number, on the right is a quantity of the denomination grams. ergs). We must, therefore, correspondingly alter the dimensions of k_0 in (9): k_0 must have the dimensions

* Cf. A. Sommerfeld and J. Runge, Ann. d. Phys., 35, 290 (1911).

§ 1. Schrödinger's Wave-Equation, the de Broglie Wave-Length 5

of the reciprocal of an action, whereas it was originally the reciprocal of a length. Since, further, k_0 must be universal, a possible value* is

$$k_0 = \frac{2\pi}{h} \quad . \quad . \quad . \quad (10)$$

where, of course, the factor 2π is arbitrary *a priori*, but *a posteriori* can be justified without ambiguity by comparison with known solutions in the quantum theory. We therefore obtain from (9) as *the final differential equation of micro-mechanics in its simplest form* (that is, for a single point-mass and a conservative field of force)

$$\Delta\psi + 2m(E - V)\left(\frac{2\pi}{h}\right)^2\psi = 0 \quad . \quad . \quad . \quad (11)$$

We call this equation the *wave-equation*,† and ψ the *wave-function*, and we regard eqn. (11) as the foundation of *wave-mechanics*.

Firstly, we shall broaden this foundation. Instead of one point-mass we shall consider several freely moving point-masses, which are coupled to each other by conservative forces. Thus V is now to be a function of the co-ordinates of position of these various point-masses. The case in which the forces have no potential (magnetic field) will be deferred, until we can proceed with the relativistic generalisation of our wave-equation (cf. § 9).

If we trace the derivation of (11) backwards, we find that in the energy law the kinetic energies must be superimposed on each other in the form

$$\sum_{(\alpha)} \frac{1}{2m_\alpha} (p_x^2 + p_y^2 + p_z^2)_\alpha$$

where the index α labels the individual point-masses. If we express the p 's by means of S and pass from S to ψ , we get as the generalisation of (11), after taking the factor k_0 from (10) over to the left side,

$$\sum_{\alpha} \frac{h^2}{8\pi^2 m_\alpha} \Delta_\alpha \psi + (E - V)\psi = 0 \quad . \quad . \quad . \quad (12)$$

The index α attached to Δ denotes, as is evident, that the differential parameter Δ is to be formed for the Cartesian co-ordinates $x_\alpha, y_\alpha, z_\alpha$ of the point-mass α . ψ itself is a function of all these co-ordinates, and cannot in general be separated into parts each of which depend only on the co-ordinates of *one* point.

* It is to be regretted that when Planck introduced his constants he did not call his element of energy $h\nu$ instead of $h\nu$. In the latter case, on account of $\omega = 2\pi\nu$, the value of h would have been equal to $\frac{1}{2\pi} 6.55 \cdot 10^{-27}$ erg sec. In place of (10) we should then have had $k_0 = 1/h$, and all other formulae of wave-mechanics would likewise have become simpler. In the sequel we shall occasionally make use of the abbreviation k_0 in (10) in order to get rid of the inconvenient factor 2π .

† Schrödinger himself originally wished to reserve the name "wave-equation" for one analogous to (5) but containing the time. We shall call the latter, to distinguish it from (11), the "time-equation." It is, as we shall see in § 5, differently built up from (5).

There is no difficulty in introducing curvilinear co-ordinates for one or several point-masses in place of the Cartesian co-ordinates. It is only necessary to transform the differential expressions Δ_x in the wave-equation into the new co-ordinates by the well-known rules. If the point-masses are not free but *bound*, their co-ordinates being subject to equations of constraint, then we must introduce, as in ordinary mechanics, generalised (Lagrange) co-ordinates, by which the equations of constraint may be eliminated. But we prefer to discuss these general matters later when we deal with a particular example (§ 11).

Schrödinger has also developed a very interesting method of deriving the wave-equation for the special case of a point-mass or for the general case of a mechanical system from a *problem in the calculus of variations*. We shall, however, also deal with this matter later in connection with the relativistic generalisation of the wave-equation.

As for the integration of the wave-equation we shall here make only the following remarks. The objective is to find such integrals of ψ as are *one-valued and continuous in the whole range of validity of the co-ordinates*, including the *boundary points*. The *boundary condition of continuity* that presents itself in this way furnishes us, to our surprise, with a complete substitute for our quantum condition (1). The quantum number n that occurs in (1) comes into existence spontaneously in solving the "boundary problem" in question. *In this way our quantum condition (1) may be dispensed with as a special axiom.* In many cases, moreover, as we shall see later, it is subjected to correction.

The position here is similar to that in the boundary problems of ordinary mechanics, for example, in that of the vibrating string. In this case, too, the boundary condition (given by the fixed ends of the string) leads to the introduction of a whole number n which distinguishes the different forms of vibration as fundamental tone and overtones, and is equal to the number of loops or one more than the number of nodes.

To be able also to dispense (in a certain sense) with the frequency condition (2) in the wave-mechanical treatment we must elaborate our wave-equation further (§ 5) by taking into account the dependence on time.

Finally we consider as the simplest conceivable example to which the wave-equation can be applied a *point-mass* under no forces ($V = 0$). According to (11) we then have the differential equation

$$\Delta\psi + k^2\psi = 0, \quad k^2 = \frac{8\pi^2}{h^2}mE \quad . \quad . \quad . \quad (13)$$

We integrate it as in the optical problem of the plane wave. Choosing the positive direction of the x -axis we write

$$\psi = Ae^{ikx} \quad . \quad . \quad . \quad . \quad (14)$$

The range of the x -co-ordinate extends from $x = -\infty$ to $x = +\infty$. Our solution is one-valued and continuous, including these limiting points, and so satisfies our general boundary condition for every positive value of E . While in other cases the boundary condition can be fulfilled only by the special choice of E , E here remains undetermined. We interpret E as the kinetic energy of our point-mass, setting $E = mv^2/2$. From (13) we then get

$$k = \frac{2\pi mv}{h} \quad . \quad . \quad . \quad . \quad . \quad (15)$$

The spatial period of our ψ -function, that is, the wave-length λ therefore becomes

$$\lambda = \frac{2\pi}{k} = \frac{h}{mv} \quad . \quad . \quad . \quad . \quad . \quad (16)$$

In this way a wave-length λ is allocated to the motion of translation of a point-mass. We call (14) a *de Broglie wave* and (16) the *de Broglie wave-length*, for Louis de Broglie—in his Thesis (Paris, 1924)—assigned wave-lengths in this way even before Schrödinger's papers appeared. The train of ideas which led him to do this will be indicated in § 5, where we shall also see the part played by the velocity v of the mass-particle (group-velocity as compared with phase-velocity) in the wave-picture of the ψ -function. At this stage we shall give no details but merely emphasise the fact that the conception of corpuscles (electrons and atoms) as waves has already borne fruit experimentally. (Details are given in Chap. II, §§ 5 and 6.) We already speak of a "diffraction of electrons by crystal lattices," and compare the experimentally obtained directions of diffraction with those which occur in the case of Röntgen rays. The twofold nature of light as a light-wave and as a light-quantum is thus extended to electrons and, further, to atoms: their wave-nature is asserting itself more and more, theoretically and experimentally, as concurrent with their corpuscular nature.

§ 2. Explanation of the Mathematical Method. Spherical Harmonics. Bessel Functions

In the sequel we shall continually encounter linear differential equations which are to be integrated in such a way that the solutions are single (uniform, *eindeutig*) within a prescribed region and, including the boundary points, continuous. As a rule this is possible only if there occurs in the differential equation a disposable parameter, to which appropriate values may be assigned. These values are called *proper* values* and the corresponding solutions are called *proper functions*. The

* *Eigenwerte*, which is often rendered by *characteristic values*; likewise *Eigenfunktionen* = characteristic functions. We have preferred to use the shorter word "proper," especially as it seems to have established itself in the current literature on the subject.

whole theory was originally developed for a vibrating string with a mass distribution varying from point to point or, what comes to the same thing, for the conduction of heat along a rod in which the conductivity varies along its length (Sturm-Liouville problems).^{*} We here give only so much of the theory as is absolutely necessary and immediately illustrate the process by taking special examples.

A. Spherical Harmonics

Starting from the vibration-equation in three dimensions

$$\Delta u + k^2 u = 0,$$

such as occurs, for example, in acoustics, let us consider a problem of boundary values with spherical symmetry and introduce polar co-ordinates r, θ, ϕ . As is well known, we obtain

$$\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial u}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial u}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 u}{\partial \phi^2} + k^2 u = 0. \quad (1)$$

We wish to integrate this equation by the method of separation of variables (cf. the corresponding method for Hamilton's differential equation in ordinary mechanics, I, p. 559) that is, we set

$$u = R(r)\Theta(\theta)\Phi(\phi).$$

Whereas R is fully determined only by a boundary condition which would have to be prescribed in the case of an external bounding sphere, Θ and Φ are already defined by the postulate that the solution is to be uniform (single-valued and continuous, *eindeutig*) in the co-ordinate regions

$$0 \leq \theta \leq \pi, \quad -\pi \leq \phi \leq +\pi$$

except for a disposable whole number in each case, as we shall presently see.

We first consider $\Phi(\phi)$. We may call ϕ a *cyclic* co-ordinate (cf. I, pp. 443 and 561) since it does not occur explicitly in the equation (1). It is in accordance with this that we set $\Phi(\phi) = e^{\pm im\phi}$. The postulate of uniformity leads to integral values of m .

We insert Φ in the differential equation, divide by $R\Theta\Phi$ and multiply by r^2 . This gives

$$r^2 \left(\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + k^2 R \right) = - \frac{1}{\Theta} \left(\frac{1}{\sin \theta} \frac{d}{d\theta} \left\{ \sin \theta \frac{d\Theta}{d\theta} \right\} - \frac{m^2 \Theta}{\sin^2 \theta} \right). \quad (1a)$$

The common value of the two sides of this equation must equal a constant, say λ . Hence we get the differential equation for Θ

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left\{ \sin \theta \frac{d\Theta}{d\theta} \right\} + \left(\lambda - \frac{m^2}{\sin^2 \theta} \right) \Theta = 0 \quad (1b)$$

^{*} For this and for all other mathematical problems involving boundary values consult the excellent volume by Courant and Hilbert, *Methoden der mathematischen Physik*, Springer, 1924.

The "separation-constant" λ is at the same time the "proper-value parameter" of this differential equation. We introduce $x = \cos \theta$ as an independent variable, write $\Theta(\theta) = y(x)$ and take into account that

$$\sin \theta \, d\theta = -dx, \quad \sin \theta \frac{d\Theta}{d\theta} = -(1-x^2)y'.$$

From (1b) there then results

$$(1-x^2)y'' - 2xy' + \left(\lambda - \frac{m^2}{1-x^2}\right)y = 0 \quad (2)$$

This is the differential equation of the *general* (so-called associated) Legendre function. The product $y(\cos \theta)\Phi(\phi)$ is a "surface-harmonic," being a so-called "tesseral harmonic."*

Singular points of a linear differential equation is the name given to values of the independent variable x for which one of the coefficients becomes infinite. It is implied that the equation is reduced to its standard form in which the coefficient of the highest derivative of y is reduced by division to unity, or we shall say the equation has been *normalised*. All other points are called *ordinary*. A singular point is called a *pole* (*ausserwesentlich singuläre Stelle*) if a power series is possible at it which contains only a finite number of members with negative exponents. Otherwise the singularity is an *essential singularity*. Ordinary points and poles are also classed together under the name of *regular points*. To determine the exponent α of the initial term of the power series at a point $z = x - x_0 = 0$, we make the substitution

$$y = z^\alpha(a_0 + a_1z + a_2z^2 + \dots) \quad (3)$$

and determine α as the root of a quadratic equation (we restrict ourselves to differential equations of the second order, which we assume normalised in the above sense) namely, of the so-called *characteristic equation*. It is obtained if the series (3) is inserted in the left-hand side of the differential equation and if the factor of the lowest power $z^{\alpha-2}$ is set equal to zero. Equating the factors of the following higher powers to zero gives us a recurrence formula for the a_k 's. The recurrence formula may be obtained more conveniently if we change the dependent variable by setting

$$y = z^\alpha v, \quad v = \sum a_n z^n \quad (4)$$

and calculate the a_k 's from the differential equation for v .

The general criterion for a pole is, as the process just described immediately shows, that the coefficients of y'' , y' and y in the original differential equation may not at the point in question, $z = 0$, approach infinity relatively to each other more rapidly than

$$1, \quad \frac{1}{z}, \quad \frac{1}{z^2}, \quad \dots \quad (5)$$

respectively.

If the two roots α_1, α_2 of the characteristic equation differ by a whole number, peculiarities (logarithmic terms) occur in one of the two

* The terms zonal, tesseral and (for the special case of maximum values of m) sectorial surface harmonics are due to Maxwell. Cf. the important Chapter IX of his *Treatise on Electricity and Magnetism*.

Particular Solutions. We shall not discuss this aspect as we are interested only in the continuous solutions of our differential equations.

In the case of the differential equation (2) the points $x = \pm 1$ and $x = \infty$ are singular, being, indeed, as we shall show, poles. If, for example, to investigate the point $x = 1$, we set $z = x - 1$, then (2) becomes

$$y'' + \frac{2}{z} \frac{z+1}{z+2} y' - \left(\frac{\lambda}{z(z+2)} + \frac{m^2}{z^2(z+2)^2} \right) y = 0.$$

The application of criterion (5) shows us that $z = 0$ is in truth a pole. If we use substitution (3) and form the factor of $z^{\alpha-2}$ the expression

$$\left(\alpha(\alpha-1) + \alpha - \frac{m^2}{4} \right) a_0$$

follows. Since we may assume a_0 to differ from zero, the characteristic equation becomes

$$\alpha^2 - \frac{m^2}{4} = 0, \quad \alpha = \pm \frac{m}{2} \quad . \quad . \quad . \quad (6)$$

The same exponents result for the point $x = -1$ by making the substitution $z = x + 1$.

To investigate the point $x = \infty$ we make the substitution $x = \frac{1}{t}$, well known from the Theory of Functions, and obtain in the case of (2):

$$\dot{y} + \frac{2t}{t^2-1} \dot{y} + \frac{1}{t^2(t^2-1)} \left(\lambda - \frac{m^2 t^2}{t^2-1} \right) y = 0$$

where the dots denote differentiations with respect to t .

The application of the criterion (5) shows as above that $t = 0$ is also a pole. The characteristic equation becomes

$$\alpha(\alpha-1) - \lambda = 0.$$

On account of the postulate that y must be continuous in the region $-1 \leq x \leq +1$ we now look for that branch of the function that has the exponent $+\frac{m}{2}$ (m being assumed positive) at the two limits $x = \pm 1$.

As suggested by eqn. (4) we therefore set

$$y = (1-x^2)^{\frac{m}{2}} v \quad . \quad . \quad . \quad (7)$$

by simultaneously detaching from y the *two* characteristic powers x^{α} of the singularities situated in finite regions ($\alpha = +m/2$, $z = x \mp 1$). By working out (2) we easily get for v the differential equation

$$(1-x^2)v'' - 2(m+1)xv' + (\lambda - m - m^2)v = 0 \quad . \quad . \quad (8)$$

This may be integrated by assuming

$$v = \sum a_r x^r \quad . \quad . \quad . \quad (9)$$

By substituting in the differential equation and equating the factor of x^v to zero we get the following recurrence formula for the a_v 's:

$$(v+2)(v+1)a_{v+2} = \{v(v-1) + 2(m+1)v - \lambda + m + m^2\}a_v \quad (10)$$

If we choose λ so that the factor of a_v vanishes, say, for $v = k$, then according to the recurrence formula all coefficients a_{k+2} , a_{k+4} , . . . vanish, and our series (9) ceases at the term $v = k$: it contains, as we also see from the recurrence formula, only even or odd powers of x according as we allow it to begin with a_0 or a_1x . The choice of λ thus decides the proper value of the spherical functions. With $v = k$ we get

$$\begin{aligned} \lambda &= k(k-1) + 2(m+1)k + m(m+1) \\ &= (k+m)(k+m+1). \end{aligned}$$

As v in this way becomes of degree k we get $k+m$ for the degree of the corresponding proper function y (cf. 7); we write n for $k+m$. Written in terms of this n the proper value is simply

$$\lambda = n(n+1) \quad (11)$$

The above method of determining a proper value and proper function is applicable in all cases in which the differential equation leads to a recurrence formula with two terms. We shall see that this is so with all quantum problems that are exactly soluble (that do not require methods of approximation or perturbations). The fact that the recurrence formula has two terms may be recognised directly from the differential equation. For if we substitute for v a power x^v then, after multiplication with the coefficients of the differential equation, only two powers may occur (in our case they are x^v and x^{v-2}). It may also be determined in general whether a given differential equation can be transformed into one having a recurrence formula with two members.*

This procedure, which is based on the termination of a power series, we shall call the *polynomial method*. It is evident that it is sufficient in character. For our polynomial (even after being multiplied by the detached factors, cf. (7)) is certainly continuous in the region under consideration (which is here $x = -1$ to $x = +1$), and so represents a proper function. The fact that our procedure is also necessary in character, that is, that there are no other proper functions of the differential equation besides those found, may easily be shown,† at any rate in the case of spherical harmonics, but may here be passed over.

The usual method of denoting the y of eqn. (7) is $P_n^m(x)$. Since k , being the degree of the polynomial (9), is a positive whole number and since m must also be given positive and integral values in the proper functions (integral on account of the postulate of uniformity for $\epsilon \pm i\mu\phi$), therefore $n = k+m$ also becomes an integer greater than or equal to m . Thus if we fix n there are $n+1$ proper functions P_n^m . The first of these, for which $m = 0$, is the spherical harmonic of n th degree in the narrower

* Forsyth, *Differential Equations*. Note on p. 589 *et seq.* of the German edition.

† Most simply by the theory of functions, cf. K. Bechert, *Ann. d. Phys.*, **83**, 906 (1927).

sense and is written simply P_n (zonal harmonic or Legendre's polynomial). The other n proper functions are the "associated" functions of the same degree n . The last proper function, for which $m = n$ (sectorial harmonics), is, by (7), proportional to

$$(1 - x^2)^{\frac{m}{2}} = (\sin \theta)^m.$$

If we differentiate eqn. (8) with respect to x , we get a differential equation for v' which is distinguished from that for v only by having $m + 1$ in place of m . From this we infer that in the series of functions $P_n, P_n^1, \dots, P_n^m, P_n^{m+1}$, the v belonging to every following polynomial may be

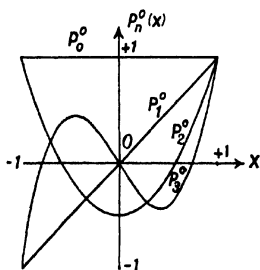


FIG. 1.

The (ordinary) spherical harmonics P_n — P_n^0 are drawn to scale:

$$P_0^0 = 1, P_1^0 = x, P_2^0 = \frac{3}{2}x^2 - \frac{1}{2}$$

$$P_3^0 = \frac{5}{2}x^3 - \frac{3}{2}x.$$

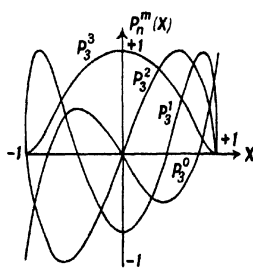


FIG. 2.

The associated spherical functions

$$P_m(x) = (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n^0(x)$$

have been divided by appropriate factors so that the maximum of each P_n^m became equal to 1. The figure shows

$$P_3^0, 1, P_3^1, \frac{\sqrt{15}}{8}, P_3^2, \frac{\sqrt{3}}{10}, P_3^3, \frac{1}{15}.$$

obtained from the preceding one by differentiation. We thus get the associated spherical harmonics represented by the Legendre polynomials

$$P_n^m(x) = (1 - x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_n(x) \quad (12)$$

In this identity the usual normalisation of the associated P_n^m 's has already been carried out. The P_n 's themselves have been normalised since Legendre's time by setting

$$P_n(1) = 1 \quad (12a)$$

But later we shall become acquainted with another kind of normalisation based on the orthogonal relationships that subsist between all proper functions.

We have yet to mention that the number of spherical harmonics

(proper functions of the two-dimensional problem in θ and ϕ) is not equal to $n + 1$, but, on account of the double sign in (12b), to $2n + 1$. For they are represented in their totality by

$$u_{nm} = P_n^m(\cos \theta) e^{\pm im\phi} \quad (12b)$$

B. Bessel Functions

We revert to eqn. (1) and now consider the radial component R of the solution there assumed. By (1a) its differential equation is

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left(k^2 - \frac{\lambda^2}{r^2} \right) R = 0 \quad (13)$$

or, if we substitute for λ the proper value (11) and set $kr = \rho$ (dashes here as well as in the sequel denote derivatives with respect to ρ):

$$R'' + \frac{2}{\rho} R' + \left(1 - \frac{n(n+1)}{\rho^2} \right) R = 0 \quad (13a)$$

According to the meaning of r the boundary points of the region in ρ are the values $\rho = 0$ and $\rho = \infty$. The criterion (5) tells us directly that $\rho = 0$ is a *pole*. We investigate the *asymptotic behaviour* at this point by a method which is perhaps lacking in mathematical rigour but which, in later cases as well as here, will always lead directly and surely to the result. For great values of ρ (13a) becomes

$$R'' + R = 0 \quad (13b)$$

and is integrated, by the two particular solutions

$$R = A e^{i\rho} \quad \text{and} \quad R = B e^{-i\rho} \quad \text{respectively} \quad (14a)$$

We can immediately find a second approximation by assuming, for example, the coefficient A to be a "slowly varying quantity." By this we mean (see p. 4) that we regard A' as differing from zero, but neglect A'' and A'/ρ as well as A/ρ^2 . By substitution in (13a) we get, if we suppress the common factor $e^{i\rho}$

$$2iA' + \frac{2i}{\rho} A = 0, \quad A = \frac{\text{Const.}}{\rho} \quad (13c)$$

The same holds for B . Both particular solutions are not only finite at $\rho = \infty$ but actually vanish. From this it follows that the solution (14), which, for $\rho = \infty$ must resolve into a combination of the particular solutions (14a), fulfils the condition of continuity that is to be postulated, not only for $\rho = 0$ but also for $\rho = \infty$; infinity introduces no new condition for the proper function R . The parameter k that also occurs in the differential equation (13) thus remains undetermined for any unlimited region. If, on the other hand, the region were limited by a sphere $r = a$, for which R , say, is to vanish, we should get a transcendental equation for k and a "discontinuous spectrum" of values of k . In the case of the unlimited region we can talk of a "continuous spectrum."

14 Chapter I. Introduction to Wave-Mechanics

We next show that our proper function (14) is essentially a Bessel function, namely that

$$R = \frac{C}{\sqrt{\rho}} J_{n+\frac{1}{2}}(\rho) \quad (15)$$

If, actually, we insert this in (13a) we get for J after an easy calculation the differential equation

$$J'' + \frac{1}{\rho} J' + \left(1 - \left(n + \frac{1}{2}\right)^2 \frac{1}{\rho^2}\right) J = 0.$$

This is the well-known differential equation of the Bessel function of index $n + \frac{1}{2}$. The fact that in (15) we can be dealing only with the solution of this function which is finite in a finite region, and not with the so-called Hankel solutions H , follows from our postulate of continuity at $\rho = 0$.

Whereas in the sequel we shall continually meet with spherical harmonics, we shall not need the Bessel functions directly. But the method will in many cases be useful for obtaining asymptotic representations such as we worked out in this example.

§ 3. Oscillator and Rotator. Their Proper Values according to Wave-Mechanics

In this section we first correct I, Chap. IV, § 1, by replacing the former quantum condition by the postulate of continuity of the proper functions concerned.

A. The Linear Harmonic Oscillator

When displaced a distance x it has the potential energy

$$V = \frac{k}{2} x^2 = \frac{m}{2} \omega_0^2 x^2 \quad (1)$$

$\omega_0 = 2\pi\nu_0$ is the "circular" frequency (*Kreisfrequenz*) of its proper vibration (in the sense of classical mechanics). This is calculated from the constants k in the expression for the elastic restoring force ($-kx$) and the mass m of the oscillator: $\omega_0^2 = k/m$. By eqn. (11) of § 1 the wave-equation of the oscillator becomes, for the given value of V ,

$$\frac{d^2\psi}{dx^2} + (\lambda - \alpha^2 x^2)\psi = 0 \quad (2)$$

λ and α are abbreviations for the following quantities:

$$\lambda = \frac{8\pi^2}{h^2} mE, \quad \alpha = \frac{2\pi m \omega_0}{h} \quad (2a)$$

The range of the co-ordinate is from $x = -\infty$ to $x = +\infty$. These limiting points of the range are essential singularities. We recognise this if we enquire into the asymptotic behaviour of ψ for large values of x .

$\psi'' = \alpha^2 x^2 \psi$. This equation is then asymptotically integrated (cf. eqn. (13b) of the previous section) by

$$\psi = e^{\pm \frac{\alpha}{2} x^2} \quad (3)$$

For it gives

$$\psi' = \pm \alpha x \psi, \quad \psi'' = \alpha^2 x^2 \psi$$

(in the latter we have neglected a term which is of lower order in x). We can use only the lower sign in equation (3) since ψ is to remain finite for $x = \pm \infty$. We therefore set

$$\psi = e^{-\frac{\alpha}{2} x^2} v \quad (4)$$

(similarly to § 7 of the preceding section) and determine v (without neglecting any terms) from the differential equation (2). Writing

$F = e^{-\frac{\alpha}{2} x^2}$ we get from (4)

$$\begin{aligned} \psi' &= F(v' - \alpha x v) \\ \psi'' &= F(v'' - 2\alpha x v' - \alpha v + \alpha^2 x^2 v). \end{aligned}$$

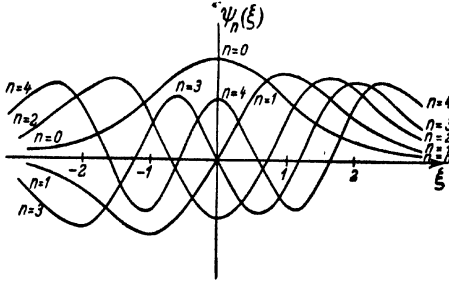


FIG. 3.

The first five Proper Functions of the oscillator

$$\psi_n = \frac{1}{2^{n/2} \sqrt{n!}} H_n(\xi) e^{-\frac{\xi^2}{2}}$$

$$H_0 = 1, H_1 = 2\xi, H_2 = 4\xi^2 - 2, H_3 = 8\xi^3 - 12\xi, H_4 = 16\xi^4 - 48\xi^2 + 12.$$

Concerning the normalisation here chosen for ψ cf. § 6 A.

So, noting that the term in x^2 cancels in virtue of our choice of the factor F , we get

$$v'' - 2\alpha x v' + (\lambda - \alpha)v = 0. \quad (5)$$

We divide by α and use as the independent variable the dimensionless quantity

$$\xi = \sqrt{\alpha} x \quad (5a)$$

Denoting differentiations with respect to ξ by dots we get from (5)

$$\ddot{v} - 2\xi \dot{v} + \left(\frac{\lambda}{\alpha} - 1\right)v = 0. \quad (6)$$

We integrate this equation by means of the power series

$$v = \sum a_\nu \xi^\nu \quad . \quad . \quad . \quad . \quad . \quad (6a)$$

and obtain from (6) by equating to zero the member with ξ^ν the two-termed recurrence formula

$$(\nu + 2)(\nu + 1)a_{\nu+2} + \left(\frac{\lambda}{\alpha} - 1 - 2\nu\right)a_\nu = 0.$$

We shall make the power series end with the term $\nu = n$. To do this we need only equate to zero the factor of a_n in the recurrence formula, whereupon all the later coefficients a_{n+2} , a_{n+4} , . . . vanish.*

Therefore

$$\frac{\lambda}{\alpha} = 2n + 1 \quad . \quad . \quad . \quad . \quad . \quad (7)$$

or, on account of (2a)

$$E_n = \left(n + \frac{1}{2}\right) h \frac{\omega_0}{2\pi} \quad . \quad . \quad . \quad . \quad . \quad (7a)$$

In the earlier treatment by means of the quantum condition $\int p dq = nh$ we obtained $nh\nu_0 = nh\omega_0/2\pi$ for the n th energy-level. Thus the characteristic difference between the old and the new formula (7a) consists in the appearance of half-integers instead of integers.

Concerning the experimental confirmation of this result, cf. the results quoted in § 3, D, in particular the footnotes (*) and (†) on page 24.

The polynomials (6a) which we get are called *Hermitean polynomials* (when appropriately normalised, which is discussed in § 6). We write $v = H_n$ and get for the n th proper function of the oscillator, by (4),

$$\psi_n = e^{-\frac{\xi^2}{2}} H_n(\xi) \quad . \quad . \quad . \quad . \quad . \quad (8)$$

The ψ_n 's are even or odd in ξ , according as n is even or odd. ψ_0 is identical with Gauss's Error Curve since $H_0 = \text{const.}$ Fig. 3† represents graphically the first five proper functions. The fact that the ψ_n 's actually fulfil the postulate of continuity in the range $-\infty \leq \xi \leq +\infty$ ensured by their expression as polynomials in (8).

B. The Rotator in Space

As in I, Chap. IV, p. 198, we imagine a point-mass m revolving round a fixed centre at a given distance a , but here we shall not assume the motion to be in a circle, that is, in one plane, but shall from the outset consider the motion to take place on a sphere, that is, we shall consider the case of two degrees of freedom. We shall return to the case of the rotator in the plane presently. For a fixed value of a the potential

* The form of the recurrence formula shows directly that in general either the even or the odd coefficients may differ from zero, but not both.

† According to Schrödinger, *Naturwissenschaften*, 14, 664 (1926).

energy is constant and may be set equal to zero. The expression $\Delta\psi$, which represents the kinetic energy in the wave-equation, is

$$\Delta\psi = \frac{1}{a^2 \sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{a^2 \sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2}$$

in three-dimensional polar co-ordinates r, θ, ϕ , by eqn. (1) in § 2, if we set $\frac{\partial}{\partial r} = 0$ and $r = a$.

Hence we have the following wave-equation with $V = 0$ and $J = ma^2$ (moment of inertia of the point-mass m about the fixed centre):

$$\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 JE}{h^2} \psi = 0 \quad (9)$$

This is the differential equation of the spherical harmonics, which were integrated in § 2 by assuming the solution $\Theta\Phi$. The proper value parameter there designated by λ is here represented by

$$\lambda = \frac{8\pi^2 JE}{h^2}.$$

Accordingly it follows from eqn. (11), § 2, that *

$$\frac{8\pi^2 JE}{h^2} = m(m+1), \quad E_m = \frac{h^2}{8\pi^2 J} m(m+1) \quad (10)$$

Compare this result with that obtained by the earlier method in I, Chapter IV, p. 198, or more directly, with I, Chapter VII, p. 414, eqn. (1). The difference is that $m(m+1)$ has taken the place of m^2 . Since $m(m+1) = (m + \frac{1}{2})^2 - \frac{1}{4}$, we see that the essential difference in this case, too, as in that of the oscillator is that *half-integers have taken the place of integers*. For if in order to apply the result to rotation-band spectra we form the difference of the proper values (energy-levels) for the transition $m_2 \rightarrow m_1$, the constant $\frac{1}{4}$ cancels out and we get

$$m_2(m_2+1) - m_1(m_1+1) = (m_2 + \frac{1}{2})^2 - (m_1 + \frac{1}{2})^2$$

in place of the $m_2^2 - m_1^2$ obtained by the earlier treatment. And we need the system of half-integral enumeration, as was remarked in § 1, if we are to account adequately for the observed rotation-bands.† The proper function

$$\psi_{m, \mu} = P_m^\mu (\cos \theta) e^{\pm i\mu\phi}$$

belongs to the proper value (10) [cf. § 2, eqn. (12b)]. In contrast with E it depends not only on m but also on μ . We have therefore not

* In place of n and m (cf. § 2) we now introduce as indices of the spherical harmonics m and μ , as we wish to reserve n for the oscillations of the rotator (cf. D).

† This holds both for the bands in the visible and for those in the infra-red region. For the latter see M. Czerny, *Zeitschr. f. Phys.*, **44**, 235 (1927); **45**, 476 (1927). The earlier explanation obtained by assuming a half-integral moment of momentum for the electrons belonging to the molecules has already been denounced as unsatisfactory in *Atombau*, Chap. IX, p. 718 (cf., in particular, footnote 2).

one proper function, but, on account of $\mu \leq m$, $2m + 1$ different ones. The proper value (10) is not simple but has a $2m + 1$ -multiplicity. In this case we speak of a degenerate proper value problem.

The notion of degeneracy was first introduced by Schwarzschild (cf. I, Note 7, p. 567). Schwarzschild used the term degenerate when rational relationships existed between the periods of the orbit, that is when the original number of "quantum degrees of freedom" allowed itself to be reduced. Bohr has often expressed the view that superfluous degrees of freedom (superfluous co-ordinates) are to be avoided and has worked out a method of doing this, for example, in the case of the Kepler ellipse and the Stark effect (cf. I, Chap. IV, § 6, and Chap. V, § 4). From the standpoint of wave-mechanics we must conversely demand that the number of degrees of freedom must from the very outset *not* be reduced. We must treat the rotator as a configuration *in space*, as we have done just above: we shall give fuller reasons for this below. We must treat the Kepler problem *not* as a one-dimensional problem ("periodic orbit" in Bohr's sense), nor as two-dimensional, but as a *three-dimensional* problem ("spatial quantising"), as was done in I, Chap. IV, § 7. Consequently, degenerate problems are the rule in wave-mechanics. Moreover, a new form of degeneracy ("exchange degeneracy," *Austausch-Entartung*) presents itself, which was unknown in the older quantum theory. The degree of degeneracy can in this case always be measured by a finite number, namely, by the *number of proper functions that belong to the same proper value*, by which we mean that non-degeneracy exists when there is only *one* proper function corresponding to the proper value in question, simple degeneracy when there are two proper functions, and so forth.

C. The Rotator in the Plane. The Oscillator in the Plane and in Space

We have the case of the rotator in the plane (motion of a point-mass in a circle), if in the expression for $\Delta\psi$ we not only make $\frac{\partial}{\partial r} = 0$, and $r = a$, but also $\frac{\partial}{\partial \theta} = 0$ and $\theta = \pi/2$, as in § 3, B; that is, we let $\Delta\psi$ depend only on the third co-ordinate ϕ . Eqn. (9) then becomes

$$\frac{d^2\psi}{d\phi^2} + \lambda\psi = 0, \quad \lambda = \frac{8\pi^2 JE}{h^2} \quad . \quad . \quad . \quad (11)$$

The solution of this equation is

$$\psi = e \pm im\phi, \quad m^2 = \lambda \quad . \quad . \quad . \quad (12)$$

where m must be an integer, in order that our solution (12) may be one-valued in ϕ . From (11) and (12) it thus follows that

$$E = \frac{h^2 m^2}{8\pi^2 J} \quad . \quad . \quad . \quad . \quad (13)$$

This agrees exactly with the earlier result of quantising the rotator in I, Chapter VII, eqn. (1), p. 414, and differs from the proper value (10) of

the rotator in space by having m^2 instead of $m(m+1)$, or, to express it less accurately, by having integral m 's in place of the half-integers $m + \frac{1}{2}$.

There can be no doubt that the treatment of the *rotator* as a *space-problem* is *exact* and that it is inadmissible to treat it as a plane problem both from the theoretical standpoint of wave-mechanics and in the light of the experimental results obtained for band spectra. This leads us to enquire whether we have also to correct our treatment of the oscillator as a *linear* configuration and whether the oscillator in the plane or in space lead to proper values other than those of the linear oscillator. To test this we replace eqn. (2) by

$$\Delta\psi + (\lambda - \alpha^2 \Sigma x_i^2)\psi = 0 \quad (14)$$

where λ and α have the same meaning as in (2a) and $\Delta\psi$ now denotes the usual differential expression in the rectangular co-ordinates x_1, x_2 or x_1, x_2, x_3 respectively. Eqn. (14) admits of "separation" into these co-ordinates just as readily as the corresponding problem in the older quantum theory (cf. I, Note 7, p. 559). If we write $\lambda = \lambda_1 + \lambda_2$ and $\lambda = \lambda_1 + \lambda_2 + \lambda_3$, respectively, we get for each co-ordinate a total differential equation of the form (2) with λ_i instead of λ and x_i instead of x , and for λ_i the condition (7) with n_i instead of n . For the sum of the λ_i 's we therefore get, by (7),

$$\lambda = \Sigma \lambda_i = \alpha \Sigma (2n_i + 1)$$

where the n_i 's are integers; and for the total energy of the oscillator we get, by (7a),

$$E = \Sigma E_i = \Sigma (n_i + \frac{1}{2}) h \frac{\omega_0}{2\pi}.$$

In the case of the plane oscillator ($i = 1, 2$), $\Sigma (n_i + \frac{1}{2})$ is certainly an integer, but in the case of the oscillator, in space ($i = 1, 2, 3$) certainly a half-integer. So we get here, just as in the case of the rotator, the surprising result that the oscillator is to be quantised in half-integers or integers alternately according to the number of dimensions used. We shall meet with the same result in the case of the Kepler problem.

For the proper functions of the oscillator in a plane or in space we simultaneously get, by (8),

$$\psi_n = \prod e^{-\frac{\alpha}{2} x_i^2} H_{n_i}(\sqrt{\alpha} x_i), \quad n = \Sigma n_i.$$

Thus the problem is degenerate: for a given n there are just as many different proper functions as the number of ways in which n may be built up additively from the integers n_i . The degeneracy disappears when we pass on to the anisotropically bound oscillator, that is, when we assume the proper frequencies ω_0 and hence also the α 's (of eqn. (2a)), for the different co-ordinate directions to be incommensurable. The

possibility of separation and the mode of calculation remains unimpaired in this case also.

Instead of rectangular co-ordinates for separating and integrating the two- or three-dimensional oscillator we may also use polar co-ordinates in two or three dimensions; cf. the fully analogous considerations in I, Note 7, page 563.

We next ask what number of dimensions is to be taken in the case of the *real* oscillator. If we are dealing with the oscillations of diatomic molecules along the line connecting the two atoms, the answer is undoubtedly the *linear* oscillator. We shall occupy ourselves further with it in § 3, D, but in conformity with the analytical nature of the problem we shall use not rectangular but polar co-ordinates. In the case of polyatomic molecules the motion becomes resolved into the principal vibrations of the system, of which each is again equivalent to a linear oscillator.

D. The Oscillating Rotator

We generalise the assumption made in § 3, B, by regarding the third co-ordinate r as variable besides the two spatial polar co-ordinates θ and ϕ . Thus we no longer set $r = a$ but introduce $\rho = \frac{r}{a}$ as a variable and imagine a field of force of potential energy $f(\rho)$ superposed, in which an oscillation of ρ can occur about the value $\rho = 1$.

The wave-equation then runs

$$\Delta\psi + \frac{8\pi^2m}{h^2}(E - f(\rho))\psi = 0 \quad (15)$$

$\Delta\psi$ is to be taken from eqn. (1) of § 2. We may write the first member of this expression (as in the classical treatment of the spherical wave)

$$\frac{1}{r} \frac{\partial^2}{\partial r^2}(r\psi) = \frac{1}{a^2\rho} \frac{\partial^2}{\partial \rho^2}(\rho\psi).$$

If we now set

$$\rho\psi = F \cdot P_m^{\mu}(\cos \theta)$$

and take into account the differential equation of the spherical harmonics [eqn. (1b) of § 2 with $\lambda = m(m+1)$], we get from (15), using, as before, the abbreviation $J = ma^2$

$$\frac{d^2F}{d\rho^2} - \frac{m(m+1)}{\rho^2}F + \frac{8\pi^2J}{h^2}(E - f(\rho))F = 0 \quad (16)$$

As suggested by I, Note 17, eqn. 5, page 612, we assume $f(\rho)$ to be of the form

$$f(\rho) = A - B\left(\frac{1}{\rho} - \frac{1}{2\rho^2} + b(\rho-1)^3 + c(\rho-1)^4 + \dots\right) \quad (17)$$

This assumption automatically fulfils the condition that $\rho = 1$ (that is, $r = a$) must be a position of equilibrium of the point-mass in the field of force and is quite general in virtue of the available constants A, B, b, c which occur in it. It corresponds to an arbitrary electrostatic field of force such as can occur between two charged ions (of which one is to be imagined at $r = 0$ and the other vibrating about $r = a$). For if we calculate the force acting in the r -direction

$$\mathbf{K} = -\frac{\partial f}{\partial r} = -\frac{1}{a}f'(\rho) = \frac{B}{a}\left(-\frac{1}{\rho^2} + \frac{1}{\rho^3} + 3b(\rho - 1)^2 + 4c(\rho - 1)^3 + \dots\right) = -\frac{B}{a}\left(\frac{\rho - 1}{\rho^3} - 3b(\rho - 1)^2 + \dots\right),$$

we find all powers of the distance $\rho - 1$ from the position of equilibrium represented. If we here set $b = c = \dots = 0$, we have almost (not quite) a *harmonic* oscillator, in which the restoring force $-\mathbf{K}$ is proportional to $r - a$ (provided we set the denominator ρ^3 equal to 1 as an approximation). The coefficient of $r - a$ is $\frac{B}{a^2}$ in this case; divided by the mass it gives the square of the frequency for small vibrations of the oscillator. If we call this frequency ω_0 as earlier, we have

$$\omega_0^2 = \frac{B}{J} \text{ or } B = J\omega_0^2 \quad (17a)$$

The assumption (17) is due to A. Kratzer. It played the decisive part in the older development of the theory of band spectra, and it offers the same advantages in the present treatment of the problem by the wave-theory.* For, as we shall presently see, it allows us to treat the rotating oscillator by the simple method of polynomials. We disregard the small correction terms with the coefficients b, c, \dots ; if we wished to take them into account we should have to supplement our method of polynomials from the theory of perturbations.

On account of (17) and (17a) our differential equation (16) becomes

$$\frac{d^2 F}{d\rho^2} + \left\{ \lambda + \alpha^2 \left(\frac{2}{\rho} - \frac{1}{\rho^2} \right) - \frac{m(m+1)}{\rho^2} \right\} F = 0 \quad (18)$$

with the abbreviations

$$\lambda = \frac{8\pi^2}{h^2} J(E - A) \quad \alpha = \frac{2\pi}{h} J\omega_0 \quad (18a)$$

We have to distinguish between two cases $\lambda > 0$ and $\lambda < 0$. We first consider the *second case* and set

$$-\lambda = \beta^2 = \frac{8\pi^2}{h^2} J(A - E) \quad (18b)$$

β , like α and λ , is a pure number.

* It was first used in this way by E. Fues, *Ann. J. Phys.*, **80**, 867 (1926); **81**, 281 (1926).

The asymptotic behaviour of F for $\rho \rightarrow \infty$ then gives from the equation (18)

$$F'' = \beta^2 F \text{ and hence } F = e^{\pm \beta \rho}$$

Since we must demand that F shall not become infinite for $\rho \rightarrow \infty$, we choose the lower sign in the exponent of e and set

$$F = e^{-\beta \rho} v \quad (18c)$$

so that

$$F' = e^{-\beta \rho}(v' - \beta v), \quad F'' = e^{-\beta \rho}(v'' - 2\beta v' + \beta^2 v).$$

From this we get, by (18), as the differential equation for the new unknown v

$$v'' - 2\beta v' + \left[\alpha^2 \left(\frac{2}{\rho} - \frac{1}{\rho^2} \right) - \frac{m(m+1)}{\rho^2} \right] v = 0 \quad (19)$$

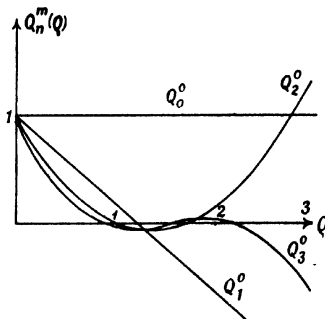


FIG. 4.

The first four polynomials Q_n^m of the oscillating rotator, etc.:

$$Q_0^0 = a_0; \quad Q_1^0 = a_0 \left(1 - \frac{\alpha^2 \cdot \rho}{\gamma(\gamma+1)} \right); \quad Q_2^0 = a_0 \left[1 - \frac{2\alpha^2 \rho}{\gamma(\gamma+2)} + \left(\frac{2\alpha^2}{\gamma+2} \right)^2 \frac{\rho^2}{2\gamma(2\gamma+1)} \right],$$

$$Q_3^0 = a_0 \left[1 - \frac{8\alpha^2 \rho}{\gamma(\gamma+3)} + \left(\frac{2\alpha^2}{\gamma+3} \right)^2 \frac{8\rho^2}{2\gamma(2\gamma+1)} - \left(\frac{2\alpha^2}{\gamma+3} \right)^3 \frac{\rho^3}{2\gamma(2\gamma+1)(2\gamma+2)} \right].$$

In drawing the figure we have chosen $a_0 = 1$, $\alpha = 10$. In reality α is mostly larger (up to $\alpha = 40$), but for such α 's the curves would approach too closely to each other. In the Q_n^m 's represented we have set $m = 0$; the Q_n^m 's for small values of m differ only slightly from the Q_n^0 's.

We integrate this by assuming

$$v = \rho^\gamma \sum a_k \rho^k \quad (19a)$$

By substitution in (19) we get for the exponent γ the characteristic equation (equating the factor of $a_0 \rho^{\gamma-2}$ to zero)

$$\gamma(\gamma-1) = m(m+1) + \alpha^2 \quad (19b)$$

that is,

$$\gamma = \frac{1}{2} \pm \sqrt{\left(m + \frac{1}{2}\right)^2 + \alpha^2} \quad (19c)$$

We must choose the upper sign in order that γ may be positive and v finite for $\rho = 0$.

The recurrence formula for the coefficients a_k (obtained by equating to zero the factor of $\rho^{\gamma+k-1}$) runs:

$$0 = [(\gamma + k + 1)(\gamma + k) - \alpha^2 - m(m + 1)]a_{k+1} + [-2\beta(\gamma + k) + 2\alpha^2]a_k \quad (19d)$$

It is a two-termed formula (*zweigliedrig*). Consequently all the a_k 's vanish for $k > n$, if we equate to zero the factor of a_k for $k = n$; v then becomes equal to the product of ρ^γ with a polynomial Q_n of the n th degree. We therefore set

$$\beta = \frac{\alpha^2}{\gamma + n} = \frac{\alpha^2}{\sqrt{(m + \frac{1}{2})^2 + \alpha^2} + n + \frac{1}{2}}; \quad (20)$$

the latter step is due to (19c).

Using (19b) and (20) we may write eqn. (19d) in the form

$$(2\gamma + k)(k + 1)a_{k+1} = 2[\beta(\gamma + k) - \alpha^2]a_k = \frac{2\alpha^2}{\gamma + n}(k - n)a_k.$$

Therefore

$$a_k = \left(\frac{-2\alpha^2}{\gamma + n} \right)^k \frac{\binom{n}{k} a_0}{2\gamma(2\gamma + 1) \dots (2\gamma + k - 1)} \quad (20a)$$

This solves our proper value problem. m is the rotation quantum, n the oscillation quantum; both occur as half-integers in our formula. According to (18b) β contains the proper value E . The polynomial component Q_n of the proper functions is graphically represented in Fig. 4 for the first values of n .

Let us now discuss eqn. (20). We observe that $\alpha \gg 1$; for $\alpha/2$ is, by eqns. (7a) and (10), equal to the ratio of the first proper value of the oscillation ($n = 0$) to the first proper value of the rotation ($m = 1$), and therefore equal to the ratio of the corresponding terms. The latter ratio may be empirically determined as the ratio of the distance between the edges to the distance between the lines in a band spectrum (cf., for example, I, Fig. 106, p. 429, or Fig. 108, p. 434), and is actually found to be a large number (of the order 20). So long as we do not assume m and n too large we may expand eqn. (20) successively as follows, retaining at present only the lowest occurring power in α :

$$\begin{aligned} \beta &= \frac{\alpha}{\sqrt{1 + \left(\frac{m + \frac{1}{2}}{\alpha}\right)^2} + \frac{n + \frac{1}{2}}{\alpha}} = \frac{\alpha}{1 + \frac{1}{2}\left(\frac{m + \frac{1}{2}}{\alpha}\right)^2 + \frac{n + \frac{1}{2}}{\alpha} + \dots} \\ &= \alpha \left[1 - \frac{1}{2}\left(\frac{m + \frac{1}{2}}{\alpha}\right)^2 - \frac{n + \frac{1}{2}}{\alpha} + \dots \right] \\ \beta^2 &= \alpha^2 \left[1 - \left(\frac{m + \frac{1}{2}}{\alpha}\right)^2 - 2\frac{n + \frac{1}{2}}{\alpha} + \dots \right], \end{aligned}$$

and hence, on account of (18a, b)

$$\begin{aligned} A - E &= \frac{J\omega_0^2}{2} \left(1 - \frac{(m + \frac{1}{2})^2 h^2}{4\pi^2 J^2 \omega_0^2} - \frac{(n + \frac{1}{2})h}{\pi J \omega_0} + \dots \right) \\ E &= \text{Const.} + \frac{(m + \frac{1}{2})^2 h^2}{8\pi^2 J} + (n + \frac{1}{2})h \frac{\omega_0}{2\pi} + \dots \\ \text{Const.} &= A - \frac{J\omega_0^2}{2}. \end{aligned}$$

So we find that, except for a constant first term, our present E is equivalent to superposing the proper value of the pure rotation (eqn. 10), on that of the pure oscillation, eqn. (7a). The constant first term is related to the work of dissociation of diatomic molecules.

We have already spoken in § 3B of the experimental confirmation of the half-integral values of the rotation quanta; in the case of oscillation quanta we have to remark that no instance is known that contradicts half-integral values. R. S. Mulliken* pointed out that a decision could be obtained by comparing the band spectra of isotopes. In the case of BO and MgH the evidence is decidedly in favour of half-integral values. In the case of other bands† (SiN, CuJ, SnCl) the decision is uncertain but seems to favour half-integral values.

But eqn. (20) gives more than the first approximation; it gives at the same time in an exact form the law according to which oscillation and rotation disturb each other in the higher quantum numbers. For example, we find without difficulty by the method given the terms of the second order to be

$$- \frac{3h^2(n + \frac{1}{2})^2}{8\pi^2 J} - \frac{3h^2(m + \frac{1}{2})^2(n + \frac{1}{2})}{16\pi^2 \omega_0 J^2} - \frac{h^4(m + \frac{1}{2})^4}{32\pi^4 \omega_0^2 J^3}.$$

The first two members agree exactly with the two correction terms of Kratzer given in I, eqn. (24), p. 616, if we set $b = c = 0$ in it, as we have done just above, with the one difference that the integral quantum numbers in the eqn. (24) mentioned are here replaced by the half-integers $m + \frac{1}{2}$ and $n + \frac{1}{2}$. The last correction member was omitted in the earlier expression as being of no great importance, but it is contained in Kratzer's original paper.‡ The close similarity between the calculations of wave-mechanics and those of the earlier quantum theory is impressively shown by this example.

For the higher values of m and n it is of course more correct not to expand eqn. (20) but to use it in its original form. For $n \rightarrow \infty$ (and also $m \rightarrow \infty$) we straightway get from (20) $\beta = 0$; thus $\lambda = 0$ and E is equal to the limiting value A , provided, indeed, that we may extrapolate our conception of a nearly harmonic attachment to this extent. The

* Phys. Rev., **25**, 259 (1925).

† Mulliken, *loc. cit.*, **28**, 1, 819 (1925); cf. also Ellis. Woldering, *Naturwiss.*, **15**, 265 (1927).

‡ *Zeitsch. f. Phys.*, **3**, 289 (1920).

limiting value would form a point of condensation both for the edges of the band ($n \rightarrow \infty$) as well as for the lines of the band of each edge ($m \rightarrow \infty$).

To define more closely the nature of the proper functions here introduced, we follow (18c) and (19a) by writing

$$F = e^{-\beta\rho} \rho^\gamma Q,$$

where Q is a polynomial of the n th degree, which is completely defined (except for a multiplying constant) by the recurrence formula (19d). By substituting $v = \rho^\gamma Q$ in the differential eqn. (19), we easily find, taking into account (19b) and (20),

$$\rho Q'' + 2(\gamma - \beta\rho)Q' + 2\beta nQ = 0.$$

If instead of ρ we introduce $x = 2\beta\rho$ as the independent variable and use dots to denote derivatives with respect to x , it follows that

$$x\dot{Q} + (2\gamma - x)\dot{Q} + nQ = 0 \quad (21a)$$

This differential equation may be compared with that of the so-called Laguerre polynomial of degree k (see § 7 *et seq.*):

$$x\ddot{y} + (1 - x)\dot{y} + ky = 0.$$

By successive differentiation of the last equation we find immediately that the i th derivative of y , which we shall call Q , satisfies the differential equation

$$x\dot{Q} + (i + 1 - x)\dot{Q} + (k - i)Q = 0 \quad (21b)$$

If we make $k - i = n$ and $i + 1 = 2\gamma$, this equation becomes (21a). Following E. Fues (*loc. cit.*) we may therefore represent our polynomial symbolically as the i th derivative of Laguerre's polynomial of degree $k = n + i$, in which, however, $i = 2\gamma - 1$ is not an integer.

We now turn to the other of the two cases to be distinguished in eqn. (18a), viz. $\lambda > 0$. By setting $\beta^2 = +\lambda$, we find asymptotically from (18) that

$$F = e^{\pm i\beta\rho}.$$

Both asymptotic solutions are here admissible, since both remain finite. Accordingly we must assume in place of (18c)

$$F = e^{-i\beta\rho} v_1 + e^{+i\beta\rho} v_2 \quad (22)$$

and in place of (19) we get the differential equation (the upper sign applying to v_1 , the lower to v_2)

$$v'' \pm 2i\beta v' + \left[\alpha^2 \left(\frac{2}{\rho} - \frac{1}{\rho^2} \right) - \frac{m(m+1)}{\rho^2} \right] v = 0.$$

Investigation of the zero point gives the same value of γ as in eqn. (19c) with the same choice of signs. We thus have one particular solution v_1 or v_2 finite at the zero point. But these solutions, which are by the way essentially identical after multiplication by $e^{\pm i\beta\rho}$, are not, as before, polynomials. The recurrence formula (19d) which holds even now if we replace $-\beta$ by $\pm i\beta$ in it, cannot be satisfied by real values

of β . Thus our power series do not cease at a particular term but are transcendental functions. At the same time β remains indefinite: in the case $\lambda = 0$ we have no discrete spectrum of proper values, such as before in the case $\lambda < 0$, but a continuous spectrum which at $\lambda = 0$ follows continuously on the limit of the discrete spectrum.

For every proper value λ formula (22) furnishes a corresponding proper function, the finite value of which is ensured at $\rho = 0$ by our choice of v_1, v_2 , whereas its finite value at $\rho = \infty$ has been proved asymptotically.

We shall come across similar conditions again when dealing with the Kepler problem: a discontinuous spectrum merging into a continuous spectrum, the first being dominant in the hydrogen series, the second more especially in the photo-electric effect. The fact that Schrödinger's theory yields both spectra simultaneously and by a single mathematical method is its particular virtue. The interpretation of the continuous spectrum according to the earlier quantum theory was given in *Atombau*, Chapter IX, § 7. We here go no further than showing that a similar interpretation is possible on the basis of wave-mechanics.

In eqn. (15) $E - f(\rho)$, which is the total energy minus the potential energy, may be claimed as analogous to the kinetic energy of the electron. If we insert $f(\rho)$ from (17), neglecting the anharmonic correction terms, we have when $\rho \rightarrow \infty$

$$E_{\text{kin}} = E - A + B\left(\frac{1}{\rho} - \frac{1}{2\rho^2}\right) \rightarrow E - A.$$

But in the case $\lambda > 0$ we have, by eqn. (18a), that $E - A > 0$. Thus the electron can (in the language of the older quantum theory) reach infinity with a finite velocity; that is, its attachment to the molecule is destroyed, and we get electron-emission.

The advantages of the above assumption (of Kratzer) for the nearly harmonic oscillator become particularly manifest if we compare it with the rather more direct assumption, also used by Schrödinger,* for the pure harmonic oscillator. In the latter we set in place of (17)

$$f(\rho) = A + \frac{1}{2}B(\rho - 1)^2$$

and we get in place of (18), with the abbreviations used earlier for λ and a

$$\frac{d^2F}{d\rho^2} + \left\{ \lambda - a^2(\rho - 1)^2 - \frac{m(m+1)}{\rho^2} \right\} F = 0.$$

Since the asymptotic behaviour is determined, as in § 3, A, by the factor

$$e^{-\frac{\alpha}{2}(\rho-1)^2},$$

we set,

$$F = e^{-\frac{\alpha}{2}(\rho-1)^2} v$$

* E. Schrödinger, Second Communication, Ann. d. Phys., 79, p. 524. Collected Papers, Quantisation as a Problem of Proper Values, Part II.

and get as the differential equation for v

$$v'' - 2\alpha(\rho - 1)v' + \left\{ \lambda - \alpha - \frac{m(m+1)}{\rho^2} \right\} v = 0.$$

With the assumption (19a) we get for the exponent γ at the point $\rho = 0$ the equation

$$\gamma(\gamma - 1) - m(m + 1) = 0,$$

an admissible solution of which is $\gamma = m + 1$. But the recurrence formula for the a_k 's no longer has *two* terms, but *three*, so that the simple polynomial method fails. For we get for the coefficient of

$$\rho^{\gamma+k-1} = \rho^{m+k}:$$

$$[(m+k+2)(m+k+1) - m(m+1)] a_{k+1} + 2\alpha(m+k+1)a_k + [\lambda - \alpha - 2\alpha(m+k)]a_{k-1} = 0.$$

Schrödinger (*loc. cit.*), and Fues (*loc. cit.*), seek to escape by introducing $\xi = \rho - 1$ and expanding the rotation term $\frac{m(m+1)}{\rho^2}$ in powers of ξ and taking up the higher powers of ξ in the theory of perturbations which is necessary not only here but also in dealing with the anharmonic oscillator.

E. Diatomic Molecules, Translation, Rotation and Oscillation

We now take a further step in the direction of approximating to the actual conditions of the molecular model, by considering two point-masses m_1, m_2 with Cartesian co-ordinates $x_1y_1z_1$, and $x_2y_2z_2$, which exert an influence on each other owing to the potential energy V of their relative positions. Then, by § 1, eqn. (12) the wave-equation runs

$$\left\{ \frac{\hbar^2}{8\pi^2m_1} \left(\frac{\partial^2\Psi}{\partial x_1^2} + \frac{\partial^2\Psi}{\partial y_1^2} + \frac{\partial^2\Psi}{\partial z_1^2} \right) + \frac{\hbar^2}{8\pi^2m_2} \left(\frac{\partial^2\Psi}{\partial x_2^2} + \frac{\partial^2\Psi}{\partial y_2^2} + \frac{\partial^2\Psi}{\partial z_2^2} \right) \right\} + (E - V)\Psi = 0. \quad (23)$$

Just as in ordinary mechanics the kinetic energy of a system can be resolved into the translational energy of the centre of inertia and into the energy of motion relative to the centre of inertia, the differential expression of wave-mechanics may be separated into two parts by introducing the relative co-ordinates x, y, z , and the co-ordinates of the centre of inertia ξ, η, ζ . Let us define

$$\begin{array}{l|l} x = x_1 - x_2 & (m_1 + m_2)\xi = m_1x_1 + m_2x_2 \\ y = y_1 - y_2 & (m_1 + m_2)\eta = m_1y_1 + m_2y_2 \\ z = z_1 - z_2 & (m_1 + m_2)\zeta = m_1z_1 + m_2z_2 \end{array}$$

and set

$$\Psi = \psi(xyz) \cdot \chi(\xi\eta\zeta).$$

From the scheme

$$\frac{\partial}{\partial x_1} = \frac{\partial}{\partial x} + \frac{m_1}{m_1 + m_2} \frac{\partial}{\partial \xi} \quad \left| \quad \frac{\partial}{\partial x_2} = -\frac{\partial}{\partial x} + \frac{m_2}{m_1 + m_2} \frac{\partial}{\partial \xi} \right.$$

it then easily follows that

$$\frac{1}{m_1} \frac{\partial^2 \Psi}{\partial x_1^2} = \frac{1}{m_1} \frac{\partial^2 \psi}{\partial x^2} \chi + \frac{2}{m_1 + m_2} \frac{\partial \psi}{\partial x} \frac{\partial \chi}{\partial \xi} + \frac{m_1}{(m_1 + m_2)^2} \Psi \frac{\partial^2 \chi}{\partial \xi^2}$$

$$\frac{1}{m_2} \frac{\partial^2 \psi}{\partial x_2^2} = \frac{1}{m_2} \frac{\partial^2 \psi}{\partial x^2} \chi - \frac{2}{m_1 + m_2} \frac{\partial \psi}{\partial x} \frac{\partial \chi}{\partial \xi} + \frac{m_2}{(m_1 + m_2)^2} \Psi \frac{\partial^2 \chi}{\partial \xi^2}$$

In forming the sum the middle terms on the right cancel out and we get, using the abbreviation [cf. I, Chap. IV, § 4, eqn. (5)]:

$$\frac{1}{\mu} = \frac{1}{m_1} + \frac{1}{m_2}, \quad \mu = \frac{m_1 m_2}{m_1 + m_2} \quad (24)$$

in place of (23):

$$\left. \begin{aligned} \frac{1}{\mu} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) \chi + \frac{1}{m_1 + m_2} \left(\frac{\partial^2 \chi}{\partial \xi^2} + \frac{\partial^2 \chi}{\partial \eta^2} + \frac{\partial^2 \chi}{\partial \zeta^2} \right) \psi \\ + \frac{8\pi^2}{h^2} [E - V(xyz)] \psi \chi = 0. \end{aligned} \right\} \quad (25)$$

The separation into the two components ψ and χ can now be performed without difficulty since, as indicated, V is to depend only on the relative co-ordinates xyz . For if we imagine eqn. (25) divided by $\psi\chi$, only the middle term would contain the co-ordinates ξ, η, ζ . This term must therefore equal a constant. If we call this constant $-k^2/(m_1 + m_2)$ and if we define another constant E_t (energy of translation) by means of the equation

$$E_t = \frac{h^2}{8\pi^2} \frac{k^2}{m_1 + m_2},$$

we get the two equations.

$$\frac{\partial^2 \chi}{\partial \xi^2} + \frac{\partial^2 \chi}{\partial \eta^2} + \frac{\partial^2 \chi}{\partial \zeta^2} + k^2 \chi = 0 \quad (26)$$

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{8\pi^2 \mu}{h^2} (E - E_t - V) \psi = 0 \quad (27)$$

The second of these equations agrees essentially with eqn. (15), except that the "resultant mass" μ has taken the place of m and $E - E_t$, that is, the energy remaining for rotation and oscillation naturally replaces the former E . We therefore no longer need to concern ourselves with (27), as its proper values and proper functions have already been treated under D.

But eqn. (26) is also familiar to us. It is the wave-equation of the individual point-mass (here the centre of inertia) moving under no forces. As we saw in § 1, it has a continuous spectrum of proper values: all positive values of k , that is, all possible velocities are admissible. If we take the x -axis as the direction of the velocity, the corresponding proper functions are given by $e \pm ikx$.

The last remarks apply, however, only to the case where the molecule happens to be in unlimited space. If it is moving in the interior of a cavity (*Hohlraum*), restoring forces occur on its upper surface, whose potential energy would have to be taken into consideration in (26). The

spectrum of proper values (of the velocities) would, in principle, pass over into a discrete spectrum, although in the case of a large enclosure it would be very close; the nature of the proper functions would also have to be essentially modified.

§ 4. The Oscillator according to Quantum Mechanics

Even before the first Schrödinger paper appeared W. Heisenberg devised a wonderful method—we follow Heisenberg* in calling it that of quantum mechanics—which leads to the same results as wave-mechanics and goes even beyond it, since it gives not only the proper values or energy values, but also the Rules of Selection and Polarisation and, generally speaking, furnishes an unequivocal method for answering all questions relating to intensity. Through this wave-mechanics has received a powerful impulse and has, as we shall see in the next section, been able to translate these more extended results into its own language. For this reason, as well as on account of the great theoretical interest of the subject, we must next occupy ourselves with the ideas of quantum mechanics. As the simplest example we here again choose the problem of the linear harmonic oscillator.

Let the position co-ordinate of the oscillator be q , its momentum co-ordinate p . Its possible stationary states, which are known to us from spectroscopy, form a *discrete* series and are characterised by certain values of the parameters denoted by . . . E_i . . . E_k , which have the dimensions of energy, and, as we shall see, can be uniquely allocated to the classically defined energy of the oscillator. Whereas these E 's form a single discrete series, the q 's and p 's are quantities bearing two indices (q_{ik} , p_{ik}), for they are to correspond to the transition from the i th to the k th state. The radiation emitted in this transition is, we assume, monochromatic and follows the law

$$h\nu_{ik} = E_i - E_k (1)$$

for which we may write

$$\omega_{ik} = \frac{2\pi}{h}(E_i - E_k) (1a)$$

where $\omega_{ik} = 2\pi\nu_{ik}$.

Equation (1) is Bohr's frequency condition. It is not, properly speaking, derived here but is implicit in the fundamental definition of the quantities defining state (phase quantities). Empirically it has its source in the fundamental law of spectroscopy, the Combination Principle, which finds mathematical expression in the difference-relation (1).

* W. Heisenberg, Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen. Zeitschr. f. Phys., **33**, 879 (1925); Born and Jordan, Zur Quantenmechanik, *loc. cit.*, **34**, 858 (1926); Born, Heisenberg, and Jordan, Zuff Quantenmechanik, *loc. cit.*, **35**, 557 (1926). See also *The New Quantum Mechanics*, Birtwistle (Cambridge University Press).

Corresponding to the fact that the radiation ω_{ik} in the transition $i \rightarrow k$ is monochromatic the q 's and p 's are assumed in the form

$$q_{ik} = a_{ik} e^{j\omega_{ik}t}, \quad p_{ik} = b_{ik} e^{j\omega_{ik}t} \quad (2)$$

(in this whole section j denotes $\sqrt{-1}$).

Whereas the time t is still treated as a continuous variable, so that we can differentiate with respect to it, the p 's, q 's and all other phase parameters (or state parameters, *Zustandsgrößen*) appear as discrete quantities, which conforms with the nature of the quantum theory as a theory of discontinuities. We have already mentioned the "network" (cf. the conclusion of I, Chap. IV, § 1, p. 202) of states possible according to the quantum theory, and the sizes of the meshes of this network, which are given by Planck's constant h , but we made our calculations on a continuous basis, since we defined our phase quantities not only for the points of the net but also for the intervening points. In quantum mechanics, on the other hand, in the form in which it was originally conceived by Heisenberg and his collaborators, only the intersections of the net are considered and distinguished by the integers i and k . The p 's and q 's are defined for them alone. In the case of the oscillator we have to do with a two-dimensional net; in a problem involving several degrees of freedom we are concerned at first with a net of more than two dimensions. In mathematics a network of numbers of this kind, which we are acquainted with in determinants, is called a *matrix*. The "calculus of matrices" called into action by Heisenberg, and elaborated by Born, Heisenberg and Jordan for the purpose of dealing with quantum problems, discloses itself in the light of the above remarks as a method of calculation appropriate to the nature of quantum problems.

To indicate the relationship between Heisenberg's method and those of the earlier quantum theory we call to mind the Fourier expansion of the state- or phase-co-ordinate q in the i th quantum state. In general this expansion consists of a series of coefficients a_{ik} , which does not break off at any point; these a_{ik} 's are analogous to the a_{ik} 's that occur in (2). In accordance with the Correspondence Principle these Fourier coefficients were formerly used as the best available means of calculating intensities. We shall see that our present a_{ik} 's serve this purpose perfectly, so that we may regard the method of matrices as a refinement of the classical theory of correspondence. Naturally, the analogy with the Fourier series is incomplete in many respects, for example, in the manner in which the a_{ik} 's are calculated.

Consider the matrix of the q 's (eqn. (2)) a little more closely. The a_{ik} 's are complex quantities which, as in optics, embrace the amplitude and the phase

$$\begin{cases} a_{ik} = |a_{ik}| e^{j\beta_{ik}} \\ q_{ik} = |q_{ik}| e^{j(\omega_{ik}t + \beta_{ik})} \end{cases} \quad (2a)$$

By (1a) we have the following relations between the ω_{ik} 's

$$\omega_{ki} = -\omega_{ik} \quad (3)$$

If we desire the same behaviour for the phases β_{ik} , that is *change of sign* when i and k are exchanged, we must have

$$a_{ki} = a_{ik}^*, \quad q_{ki} = q_{ik}^* \quad (3a)$$

whereas we assume the opposite behaviour, that is *equality of sign* for the amplitudes which by their very meaning are positive. The stars in (3a) denote the transition to conjugate imaginaries. That is, the matrix of the q 's (and likewise that of all other phase quantities) contains conjugate imaginaries at points symmetrical to the diagonal; it is, to use the customary term, a *Hermitean matrix*.

We have now to define the rules for calculating with matrices. It is clear how this is to be done for addition and subtraction. The matrix $a \pm b$, for example, is formed by adding or subtracting all corresponding members a_{ik} and b_{ik} . It is also clear how a matrix a is multiplied by an ordinary numerical factor α . Each individual term a_{ik} has to be multiplied by α . But what is the product of two matrices to signify, for example, the product of the two matrices p and q represented in (2)? We take our cue from the rule for multiplying determinants and define

$$(pq)_{ik} = \sum_l p_{il} q_{lk} \quad (4)$$

The appropriateness of this definition becomes clear if we reflect that the complex of vibrations given by the ω_{ik} 's is fixed by the nature of our quantum mechanical problem. Whatever the calculations we may undertake we must remain within the realm of this complex of vibrations. If we insert (2) in (4), we get

$$(pq)_{ik} = \sum_l b_{il} a_{lk} e^{i(\omega_{il} + \omega_{lk})t}$$

But by (1a) we have

$$\omega_{il} + \omega_{lk} = (E_l - E_i + E_l - E_k) \frac{2\pi}{h} = \omega_{ik}$$

hence

$$(pq)_{ik} = \left(\sum_l b_{il} a_{lk} \right) e^{i\omega_{ik}t}$$

Thus the product pq , in virtue of our multiplication rule (4), belongs to the same group of vibration states as p and q individually.

But our multiplication rule leads to a further result: *we no longer have* $pq = qp$! For eqn. (4) states that

$$(qp)_{ik} = \sum_l q_{il} p_{lk} \neq (pq)_{ik} \quad (4a)$$

As, however, in ordinary calculations, the commutative law is continually being used, its absence here in calculating with matrices produces a

hiatus. This can and, indeed, *must* be bridged over by a new convention,* which is given by the "Commutation Law" (*Vertauschungs-Relation*)

$$pq - qp = \frac{h}{2\pi i} \delta \quad . \quad . \quad . \quad . \quad . \quad (5)$$

where δ is the "unit matrix" defined by

$$\delta_{ik} = \begin{cases} 1, & \text{if } i = k \\ 0, & \text{if } i \neq k \end{cases} \quad . \quad . \quad . \quad . \quad . \quad (5a)$$

But the postulate (5) is to hold in each case only for two canonically conjugate quantities (cf. I, note 4, p. 545, eqn. 20c), that is, besides for p and q themselves, also for any two quantities which are connected with the variables p and q by a canonical transformation (*ibid.* p. 546). On the other hand, in the case of two quantities of the character of the q 's or two of the character of the p 's the right-hand side of (5) must be replaced by zero: so we may say that two quantities of the same kind may be interchanged or are commutative.†

These few rules are already sufficient to make the problem of the oscillator accessible to the calculus of the matrices. To deal with other problems and to formulate quantum mechanics generally it is naturally necessary to go much further, for example, to differentiate one matrix with respect to another, questions into which we shall not enter here.

In the following calculations we shall take our stand on the relations

$$\ddot{q} + \omega_0^2 q = 0 \quad (6) \qquad p = m\dot{q} \quad (7) \qquad H = \frac{1}{2m}(p^2 + m^2\omega_0^2 q^2) \quad (8)$$

which we take directly from classical mechanics.

The first of these is the classical differential equation of the oscillator, the second the usual definition of momentum, the third the classical expression for the energy as a function of the p 's and q 's (H denotes "Hamilton's Function," as in I, Chap. IV, p. 194). As in eqn. (1) of the preceding section ω_0 is the classical proper frequency.

This close link with classical mechanics is characteristic of quantum mechanics. It is not the axioms but the methods of calculation that are altered and adapted to the discontinuous character of quantum problems. The following discussion will show that these methods of calculation, however strange they may appear at first sight, are *au fond* quite elementary and lead by the most direct route to the object in view.

Eqn. (6) written down for a single element ik runs, if we insert q from (2),

$$(\omega_0^2 - \omega_{ik}^2)q_{ik} = 0.$$

*In this convention Heisenberg was originally guided by a "summation law" of Thomas and Kuhn, cf. Chap. II, § 3, C.

†An introduction to the calculus of matrices is given by M. Born, *Probleme der Atomdynamik*. See also Birtwistle, *The New Quantum Mechanics* (Camb. Univ. Press).

From this we conclude that

$$\text{either } q_{ik} = 0 \text{ or } \omega_{ik} = \pm \omega_0;$$

or, expressed differently: *all the q_{ik} 's vanish with the exception of those for which $\omega_{ik} = +\omega_0$ or $-\omega_0$.*

It is convenient to distinguish these two possibilities by using a special order in the numbering of the matrix elements, as this order is still quite open to choice. We shall take

$$\begin{aligned} \omega_{ik} = +\omega_0 & \text{ corresponds to the transition } i \rightarrow i-1 \\ \omega_{ik} = -\omega_0 & \text{ " " " } i \rightarrow i+1 \end{aligned} \quad (9)$$

that is, in the former case let $k = i-1$, and in the latter $k = i+1$. We have correspondingly for the q 's

$$\begin{aligned} q_{ik} &= 0 \text{ for } k \neq i \mp 1 \\ q_{ik} &\neq 0 \text{ for } k = i \mp 1 \end{aligned} \quad (9a)$$

With this convention for the numbering of the matrix elements, then all the *diagonal elements* ($k = i$) vanish in the matrix of the q 's, and also *all the elements not adjacent to the diagonal* (for which, therefore, $|k-i| > 1$); only the *two series next to the diagonal* contain q 's that do not vanish.

Concerning eqns. (9) it must be added that they may be summarised and generalised in the form:

$$\omega_{ik} = (i-k)\omega_0. \quad (9b)$$

The values of the oblique rows adjacent to the principal diagonal in the matrix table of q 's follow from the commutative law (5). From (7) we get

$$p_{ik} = m j \omega_{ik} q_{ik}. \quad (10)$$

Using the multiplication rule (4) or (4a), respectively, we obtain from (5), with $k = i$,

$$m j \sum_l (\omega_{il} q_{il} q_{li} - q_{il} \omega_{li} q_{li}) = \frac{h}{2\pi j}$$

or, written more conveniently, with due regard to (3),

$$\sum_l \omega_{il} q_{il} q_{li} = -\frac{h}{4\pi m}. \quad (10a)$$

The summation with respect to l becomes reduced on account of (9) and (9a) to two members, namely, to those for which $l = i-1$, $\omega_{il} = +\omega_0$, and $l = i+1$, $\omega_{il} = -\omega_0$. Hence it follows from (10a) that

$$q_{i, i-1} q_{i-1, i} - q_{i, i+1} q_{i+1, i} = -\frac{h}{4\pi m \omega_0}.$$

The products on the left-hand side are by eqn. (3a) real and positive, namely, equal to the square* of the absolute value of the matrix elements in question. So, in place of the preceding equation we may write:

$$|q_{i+1, i}|^2 - |q_{i, i-1}|^2 = \frac{h}{4\pi m \omega_0} \quad (11)$$

From this we conclude that the $|q|^2$'s form an arithmetic series, which is unlimited in the positive direction but breaks off in the opposite direction as it can contain only positive terms. The indices at which we break off the series is still open to choice, as we have so far decided only upon the relative values of i , k , but not upon their absolute values. We may therefore agree to make $q_{1, 0}$ the last non-vanishing member of the series, and $q_{0, -1}$ equal to 0. With $i = 0, 1, \dots, n$ eqn. (11) then gives

$$|q_{1, 0}|^2 = \frac{h}{4\pi m \omega_0}, \quad |q_{2, 1}|^2 = \frac{2h}{4\pi m \omega_0}, \quad |q_{n, n-1}|^2 = \frac{nh}{4\pi m \omega_0} \quad (11a)$$

In the sense of eqn. (2a) we may say that this fixes the amplitude factor $|a|$ but leaves the phase β undetermined. Taking into account (3a) and (9) we write†

$$\left. \begin{aligned} q_{n, n-1} &= \sqrt{\frac{nh}{4\pi m \omega_0}} e^{i(\omega_0 t + \beta_{n, n-1})} \\ q_{n, n+1} &= \sqrt{\frac{(n+1)h}{4\pi m \omega_0}} e^{i(-\omega_0 t + \beta_{n, n+1})} \end{aligned} \right\} \quad (12)$$

We now bring in the energy equation (8). In it Hamilton's function H is itself a matrix but, as we shall see, one of specially simple construction. We call its elements H_{ik} in general; so its diagonal elements ($i=k=n$) are H_{nn} . By (10) and (4) we get, when $i=k=n$, that

$$p_n^{2n} = -m^2 \sum_i \omega_{ni} q_{ni} \omega_{in} q_{in}$$

and so, by (3) and (3a)

$$p_{nn}^2 = m^2 \sum_i \omega_{ni}^2 |q_{ni}|^2.$$

Accordingly the second summand in H becomes

$$m^2 \omega_0^2 q_{nn}^2 = m^2 \sum_i \omega_0^2 |q_{ni}|^2.$$

It therefore follows from (8) that

$$H_{nn} = \frac{m}{2} \sum_i (\omega_{ni}^2 + \omega_0^2) |q_{ni}|^2.$$

*The square of the absolute value of the matrix elements is called *Norm* in German.
† Cf. with formula (12) the formula for q_{max} which occurs in the earlier quantum theory of the oscillator and is quite similarly constructed.

§ 4. The Oscillator according to Quantum Mechanics 35

This sum also becomes reduced to two terms on account of the factor $|q_{nl}|^2$, namely to those for which $l = n \pm 1$, and, on account of (9) and (11a), yields

$$H_{nn} = m\omega_0^2 (|q_{n, n-1}|^2 + |q_{n, n+1}|^2) = (n + \frac{1}{2})\hbar\frac{\omega_0}{2\pi} \quad (13)$$

Before we discuss this remarkable result we shall supplement our previous discussion by showing that the commutative law for the q 's and the p 's is identically fulfilled in the cases $i \neq k$, and that all elements H_{ik} vanish for $i \neq k$.

For $i = n$, $k = n + \nu$, $\nu \neq 0$ the commutative law (5), by (10) and (4), requires that

$$\begin{aligned} \sum_l (\omega_{nl} q_{nl} q_{l, n+\nu} - q_{nl} \omega_{l, n+\nu} q_{l, n+\nu}) \\ = \sum_l (\omega_{nl} - \omega_{l, n+\nu}) q_{nl} q_{l, n+\nu} = 0 \quad (14) \end{aligned}$$

On account of the factor q_{nl} again only the terms $l = n - 1$ and $l = n + 1$ of the sum come into question. For $l = n - 1$, remembering that we have excluded $\nu = 0$, we find that $q_{l, n+\nu}$ differs from zero only when $\nu = -2$. In this case, however, the first factor of the sum on the right-hand side of (14) becomes, by (9), equal to

$$\omega_{n, n-1} - \omega_{n-1, n-2} = \omega_0 - \omega_0 = 0.$$

For $l = n + 1$ we get correspondingly that $q_{l, n+\nu}$ does not become zero only when $\nu = +2$ ($\nu = 0$ again being excluded). But, by (a), the first factor of the same sum runs

$$\omega_{n, n+1} - \omega_{n+1, n+2} = -\omega_0 + \omega_0 = 0.$$

Thus our postulate (14) is actually identically fulfilled for all n 's and ν 's.

Concerning the matrix elements H_{ik} , the same calculation that led to (13) gives for $i = n$, $k = n + \nu$, $\nu \neq 0$.

$$H_{n, n+\nu} = \frac{m}{2} \sum_l (\omega_0^2 - \omega_{nl} \omega_{l, n+\nu}) q_{nl} q_{l, n+\nu} \quad (15)$$

On account of the factors q_{nl} and $q_{l, n+\nu}$ all members of the sum again vanish, unless $l = n \mp 1$ and simultaneously $\nu = \mp 2$. But then the first factor vanishes. For we have in these two cases, by (9), that

$$\omega_{nl} \omega_{l, n+\nu} = \omega_{n, n \mp 1} \omega_{n \mp 1, n \mp 2} = \omega_0^2.$$

Accordingly all $H_{n, n+\nu}$'s become zero for any n 's whatsoever and $\nu \neq 0$. We may also express this by saying that H is a *diagonal matrix*.

But this simultaneously implies that *all elements of the matrix H are constant with respect to the time*. For the dependence on time, given in our expression (2) by $e^{j\omega_{ik}t}$, of course vanishes for $i = k$, since, by (1a),

ω_{ik} becomes equal to 0. Our proof that H is a diagonal matrix thus simultaneously involves *the proof of the energy law*.

Further, we must convince ourselves that the initially introduced "energy parameters" E_i , E_k , may be identified with the elements H_{ii} , H_{kk} of the energy matrix. The following verification must here suffice. By (1a) and (9b) we have

$$E_i - E_k = \omega_{ik} \frac{h}{2\pi} = (i - k) \frac{\omega_0 h}{2\pi}.$$

But, by (13), $H_{ii} - H_{kk}$ has the same value. From the differences of the E 's we may now pass on to the E 's themselves, which we normalise with respect to an additive constant (so far left undetermined) by using (13) and setting

$$E_n = H_{nn} = (n + \frac{1}{2}) \frac{\omega_0 h}{2\pi} \quad . \quad . \quad . \quad (16)$$

But this is exactly the wave-mechanical proper value of eqn. (7a) § 3. *The wave-mechanical proper values and the quantum-mechanical diagonal members of the energy matrix are identical* (in this example and also generally). In contrast, however, with the older quantum theory of the oscillator with its "whole" quantum numbers there is the repeatedly emphasised distinction of "half"-quantum numbers.

Which of the two methods is the *simpler*, that of wave-mechanics, which is analytical, or that of quantum mechanics, which is algebraic? If we leave out of account the fact that analysis is our familiar weapon, we must say that the algebraic method of quantum mechanics uses fundamentally more elementary operations; calculating with a finite number of discrete elements is easier, in principle, than calculating with a continuum. On the other hand, when dealing with problems involving several degrees of freedom the indices in the calculations with matrices swell to such a number that it is difficult to keep a clear view and the formulæ become unmanageable. In general, moreover, difficulties arise, such as have been overcome in only the most recent accounts by Hilbert and Weyl.* In actual fact the most important problems (Kepler motion, Zeeman effect, Stark effect) were first fully solved only by means of wave-mechanics. The state of affairs is much the same as in the theory of functions, where the specifically elementary methods of Weierstrass are more cumbersome than the infinitesimal methods of Cauchy and

* We are referring to the so-called "symmetrising" of the aggregates in p and q that occur in the Hamiltonian function. Since in classical mechanics, from which the Hamiltonian function was taken over, products such as pq and qp denote the same, whereas in quantum mechanics they do *not*, the method adopted in quantum mechanics is to replace pq by the symmetrical and therefore one-valued expression: $\frac{1}{2}(pq + qp)$. A corresponding method may be defined for any powers of pq and also for any arbitrary functions. In the case of the oscillator, for which the Hamiltonian function (8) contains only the squares of p and q which are, of course, in themselves symmetrical, this difficulty did not occur.

Riemann. We may perhaps express this by saying that the method of matrices has the advantage of being simpler in principle, but wave-mechanics is simpler in practice and keeps the problem under clearer survey.

We further ask: which of the two methods approaches completion the more closely? Here too we must give quantum mechanics the preference, for it has given us not only the proper values of the oscillator, but concurrently, by a *uniform* method, also the selection rules [eqn. (9) and (9a)] and the amplitudes of emission [eqn. (11a) or (12)]. This proves quantum mechanics to be a really sufficient, logically complete method of calculation and so distinguishes itself essentially from the earlier quantum theory, in which, in order to arrive at statements about intensities, we had to enlist the aid of assumptions which accorded with the correspondence principle; these assumptions had come from a realm quite different from that of quanta. The wave-mechanics that we have so far developed does not go beyond the old quantum theory, in that it furnishes us with the proper values and proper functions but says nothing about intensities. We shall, however, see in the next section that it is able to match this advantage of quantum mechanics. The correct standpoint is to recognise the rival nature of these two methods which give identical results, and to see in them a reflection of the two rival methods of optics, that of light-quanta and that of waves.

As stated at the beginning of this section the matrix q represents the *transition* from one state to another accompanied by the emission of radiation. We are, however, also deeply interested in the *states themselves*, concerning which we can get considerable knowledge by experimental means, such as, by the method of Gerlach and Stern. In classical mechanics and electrodynamics there is no antithesis between states and transitions. The oscillator radiates classically according to the measure of its *present* state. Quantum mechanics, the premises of which link up closely with classical mechanics, associates the emission with the transitions q_{ik} . For the states themselves quantum mechanics is left only with the diagonal elements q_{nn} . But these vanish by eqn. (9a). Concerning these states quantum mechanics makes only the assertion (true in itself but not sufficient) that they are radiationless. In this respect it would be at a disadvantage compared with wave-mechanics, which through the proper function provides us with a detailed description of the *states*. But we shall see in § 6 that in this respect also the two methods can exist side by side. The description of the states by means of matrices is, indeed, a little more cumbersome than that given by the proper functions, but in essence is equivalent to it.

Finally we must touch on a general epistemological point. The avowed object of Heisenberg's first paper on quantum mechanics was to develop a method "which would be based exclusively on relationships between essentially observable quantities." Ideas such as "the position of the

By setting the exponent of (2), the "phase" of the wave, equal to a constant and differentiating with respect to t , we obtain the "phase-velocity," which as in eqn. (5), § 1, we shall call a :

$$a = \frac{dx}{dt} = \frac{\omega}{k} \quad . \quad . \quad . \quad . \quad . \quad (4)$$

If we insert in this the value of ω from (1), $\omega = 2\pi\nu = 2\pi mc^2/h$ and use for the value of k that obtained from eqn. (15) of § 1, we get

$$a = \frac{c^2}{v} \quad . \quad . \quad . \quad . \quad . \quad (5)$$

Since $v < c$, it follows from this that $a > c$.

Thus the phase of the de Broglie waves is propagated with a velocity greater than that of light.

Whereas we just now referred to § 1, that is, to Schrödinger's wave-equation, de Broglie derives eqn. (5) in a very general way from a Lorentz-transformation between the reference system moving with the point-mass and the system of the observer.

We shall now show that v plays the part of a *group-velocity* for our wave-phenomenon (2). If we call this group-velocity b , we assert that

$$v = b, \quad ab = c^2 \quad . \quad . \quad . \quad . \quad . \quad (6)$$

For defining the group-velocity we have, as a counterpart to eqn. (4),

$$b = \frac{d\omega}{dk} \quad . \quad . \quad . \quad . \quad . \quad (7)$$

As group-velocity is usually explained in a rather specialised and unsatisfactory way, and as we shall come across it often in the sequel, we here give a method of deriving it based on very general assumptions. Instead of using the single wave (2) we start out from the wave-group

$$U = \int_{k_0 - \epsilon}^{k_0 + \epsilon} A(k) e^{i[kx - \omega(k)t]} dk,$$

that is, we imagine a continuous series of single waves superposed on each other and of amplitude $A dk$, in which we suppose A and ω to vary continuously with k in some way. But we shall call such a series a "group" only if the wave-numbers contained in it are sufficiently near one another; we have indicated this in the choice of the limits of integration $k_0 \pm \epsilon$. We re-write the exponent of e in the following form :

$$kx - \omega t = k_0 x - \omega_0 t + (k - k_0)x - (\omega - \omega_0)t.$$

This gives us

$$\begin{aligned} U &= C e^{i(k_0 x - \omega_0 t)}, \\ C &= \int_{k_0 - \epsilon}^{k_0 + \epsilon} A(k) e^{i\{ (k - k_0)x - (\omega - \omega_0)t \}} dk \quad . \quad . \quad . \quad (8) \end{aligned}$$

We now look for such points x, t , for which the "mean amplitude $|C|$ of the group" (not, as previously, the phase) has a constant value. Since x and t occur only in the exponent of (8) we must set

$$(k - k_0)x - (\omega - \omega_0)t = \text{const.}$$

that is,

$$\frac{dx}{dt} = \frac{\omega - \omega_0}{k - k_0}.$$

But for a sufficiently narrow group this quotient has a limiting value independent of k and depending only on k_0 , namely

$$\frac{dx}{dt} = b = \left(\frac{d\omega}{dk} \right)_{k=k_0}.$$

This proves equation (7).

b gives us the velocity with which the mean amplitude $|C|$ of the group is propagated. The propagation of the mean phase or, respectively, of the phase of the mean point k_0, ω_0 of the group is different from this and is given by eqn. (4), as before.

From our expression (7) we easily arrive by means of the relation $\omega = ka = \frac{2\pi a}{\lambda}$ at the formulæ usually given for the group-velocity, namely

$$b = a + k \frac{da}{dk} = a - \lambda \frac{da}{d\lambda} \quad . \quad . \quad . \quad (7a)$$

or to the form used by de Broglie:

$$\frac{1}{b} = \frac{d}{dv} \left(\frac{v}{a} \right) \quad . \quad . \quad . \quad . \quad (7b)$$

but formula (7) is obviously simpler and, on account of its analogy with formula (4), more instructive.

We have next to prove equation (6). By (1) we have

$$\omega = 2\pi \frac{E}{h} = \frac{2\pi}{h} \left(\dots + \frac{mv^2}{2} + \dots \right) \quad . \quad . \quad . \quad (9a)$$

The dots signify that to the term written down there is to be added a constant term which denotes the rest-mass, and also further correction terms which take into account the relativity variation of mass. On the other hand, we have, by § 1 (15) of this chapter,

$$k = 2\pi \frac{mv}{h} \quad . \quad . \quad . \quad . \quad (9b)$$

From (9a) and (9b) we form

$$d\omega = \frac{2\pi m}{h} v dv, \quad dk = \frac{2\pi m}{h} dv.$$

From this it follows that

$$b = \frac{d\omega}{dk} = v \quad . \quad . \quad . \quad . \quad (9c)$$

which was to be proved.

In this method of deriving the formula our result seems to be only an approximation for small velocities, since in eqns. (9a) and (9b) we have neglected the variation of mass with velocity. Actually, however, eqn. (9c) is correct in the form in which it is written. This is at once recognised if we follow de Broglie and use relativity from the very outset.

We take this opportunity of stating that in other respects de Broglie's original method has been abandoned in the later developments of his theory and he himself no longer maintains it.* De Broglie originally allowed the wave-function to have an infinity which served in a way as the material nucleus of the wave-system. In this manner one would lose the mathematical uniqueness which permeates Schrödinger's later theory. We regard de Broglie's theory, as does Schrödinger himself, in the light of an important precursor of wave-mechanics but not its final form.

The fact that the group-velocity b is less than the velocity c of light, becoming proportionately smaller as the phase-velocity a becomes greater, in itself shows that the dispersion of the de Broglie waves is *normal*. For, since $b < a$, eqn. (7a) states that $da/d\lambda > 0$. Consequently the inequality

$$\frac{dn}{d\lambda} < 0 \quad (10)$$

holds for the refractive index $n = c/a$ as is the case with visible light for bodies that disperse normally. The expression of $n = c/a$ as a function of the de Broglie wave-length is by eqn. (5) of this section and eqn. (16) of § 1

$$n = \frac{v}{c} = \frac{h}{mc\lambda} \quad (10a)$$

More general significance attaches to the fact that the velocity v of our point-mass does not correspond to the propagation of a monochromatic wave but to that of a wave-group. It is not the monochromatic wave that is the physical picture of the moving point-mass, but the *wave-group*, or, as we also say, the *wave-packet*.

We now pass on from the point-mass under no forces to the general problem of wave-mechanics. Here we likewise supplement the spatially defined Schrödinger ψ -function and expand it according to the assumption (3) into a space-time function which we shall call u :

$$u = \psi e^{\frac{2\pi i}{h} Et} \quad , \quad , \quad , \quad , \quad (11)$$

We here interpret E as the energy parameter that occurs in the differential equation of the ψ -function and we take E in its rationally normalised form (cf. footnote, p. 1). We have therefore to include in E not only the kinetic and potential energy of the state in question but also the rest energy (*Ruh-energie*) contained in the masses, which is represented in the case of the individual point-mass by $E_0 = m_0 c^2$. For only when E is regarded in this way is there any sense in the definition of the vibration number ν in eqn. (1). It need hardly be mentioned that the factor

* Cf. the introduction to his book on wave-mechanics published this year.

of $2\pi i$ in the exponent of eqn. (11) denotes this particular vibration number ν .

From the remarks just made about the way in which E has been normalised it follows that V must be normalised in the same way. So we may write V more clearly as

$$V = E_0 + U, \quad E_0 = m_0 c^2 \quad (12)$$

where U denotes the "energy of position" calculated in the ordinary way.

The question now arises whether we can re-write Schrödinger's equation for ψ in the form of a "time-equation." We shall have to postulate this in accordance with the optical analogy in § 1. For then the u -equation (5) which contains the time was the primary and physical datum; the spatial ψ -equation (6) arose out of it only when the monochromatic assumption (5a) had been added. But this time-equation may contain, as in the optical case, only general matter-constants and no special constants defining state. Our parameter E , which is a characteristic value (*Eigenwert*), is a state-constant (*Zustands-konstante*), since it varies in passing from one characteristic state (*Eigenzustand*) to another. It must not therefore occur in the required time-equation. Or, expressed in other words, the purpose of passing over to the time-equation is to eliminate the parameter E out of the wave-equation. We write Schrödinger's equation in the form (11) of § 1, but remark that everything that follows also applies to the problems of many electrons (of many bodies), for example, in eqn. (12) of § 1. So we start, say, from

$$-\frac{\hbar^2}{8\pi^2 m} \Delta \psi + V\psi = E\psi \quad (12a)$$

and observe that in virtue of the assumption (11) this equation is identical with the time-equation

$$-\frac{\hbar^2}{8\pi^2 m} \Delta u + Vu = \frac{\hbar}{2\pi i} \frac{\partial u}{\partial t} \quad (13)$$

We couple with (13) the conjugate equation

$$-\frac{\hbar^2}{8\pi^2 m} \Delta u^* + Vu^* = -\frac{\hbar}{2\pi i} \frac{\partial u^*}{\partial t} \quad (13a)$$

by defining u^* as

$$u^* = \psi^* e^{-\frac{2\pi i}{\hbar} E t} \quad (11a)$$

and assume ψ^* as the conjugate imaginary to ψ . Of course $\psi^* = \psi$ if, as in the example of the oscillator, ψ is real. In other cases, however (cf. the Zeemann effect, § 1), ψ is essentially complex.

This would bring us to our goal: to eliminate E out of the wave-equation and to introduce the time into this equation instead. But the result does not quite meet our expectations. For, following on from (5)

of § 1 we wished to obtain a "vibration-equation," in which the "acceleration" $\frac{\partial^2 u}{\partial t^2}$ would occur. Instead of this we have in (13) and (13a) differential equations of the type of the "diffusion equation," into which the velocities $\frac{\partial u}{\partial t}$ and $\frac{\partial u^*}{\partial t}$ enter respectively. It is true that in this case the diffusion involves imaginary "diffusion-coefficients" and hence it is not a question of an exponential decrease with the time, such as is usually characteristic of diffusion phenomena, but with a phenomenon periodic in time, such as we assumed in (11) and (11a). Conversely we may remark that a vibration-equation with imaginary propagation would entail a fading away in time, that is, would exhibit the character of diffusion.

At present eqns. (13) and (13a) are, by their derivation, fully equivalent to the eqn. (12a) which does not involve the time, and we have so far not gone a step beyond eqn. (12a) in setting them up. To make an advance we postulate that the eqns. (13) and (13a) are still to hold if the problem depends on the time in a different way from that of eqn. (11); if, for example, V also contains the time explicitly. Thus we assert that in eqns. (13) and (13a) we have the general "time-equations" of (non-relativistic) wave-mechanics, which are valid in all cases where the external forces allow themselves to be described by a potential $V(x_1, y_1, z_1 \dots t)$. They regulate not only the individual proper vibration but also, when appropriately excited, a complex of proper vibrations.

For what immediately follows we shall find the assumptions (11) and (11a) rather more fruitful than the differential eqns. (13) and (13a). For these assumptions lead us to the *completion of wave-mechanics* mentioned as a sub-title of this section.

We first recall Green's theorem in its simplest form (for an extensive generalisation we refer to § 9):

$$\int (u \Delta v - v \Delta u) d\tau = \int \left(u \frac{\partial v}{\partial n} - v \frac{\partial u}{\partial n} \right) d\sigma \quad (14)$$

where u and v are arbitrary continuous functions in the volume, denoted by $d\tau$, over which the integration is to be performed; $d\sigma$ is an element of surface of this volume. In the case of *one* point-mass $d\tau$ is a three-dimensional element of space, in the case of *several* point-masses it is a multi-dimensional element of space. If we set $u = \psi_n$, $v = \psi_m$, that is, equal to two proper functions that belong to our wave-equation (12a), and have the proper values E_n , E_m , we may allow the surface σ to move to infinity without ever encountering a singularity. The integral on the right then vanishes, so long as u and v themselves vanish with sufficient rapidity, as is always the case with discrete characteristic values. On the left-hand side Δu and Δv become proportional to $(E_n - V)\psi_n$ and

$(E_m - V)\psi_m$ respectively, and from (14) we get (the V term going out)

$$(E_m - E_n) \int \psi_n \psi_m d\tau = 0.$$

If the E 's are all different from one another and if $m \neq n$, the integral must vanish. This is the condition of orthogonality of the proper functions. For $m = n$ the integral does not, of course, vanish. We set it equal to 1 and so normalise the proper function, whose definition hitherto contained an undetermined multiplying constant. This gives us the general condition of normalisation, which differs from the occasionally mentioned special normalisations (cf., for example, eqn. (12a) in § 2). Making use of the "unit matrix" (5a) in § 4 we embrace the conditions of orthogonality and normalisation in the one formula:

$$\int \psi_n \psi_m d\tau = \delta_{nm} \quad . \quad . \quad . \quad . \quad (15)$$

If ψ is complex we shall take the reality of δ_{nm} into consideration by writing in place of (15)

$$\int \psi_n \psi_m^* d\tau = \int \psi_m \psi_n^* d\tau = \delta_{nm} \quad . \quad . \quad . \quad (15a)$$

and, in particular, for $n = m$

$$\int \psi_n \psi_n^* d\tau = 1 \quad . \quad . \quad . \quad . \quad (15b)$$

By (11) and (11a) the relation

$$\int u_n u_n^* d\tau = 1 \quad . \quad . \quad . \quad . \quad (15c)$$

then also holds.

It is at this point that the essential physical hypothesis of Schrödinger's theory enters, which first allows a comparison between theory and experiment: *the positive quantity*

$$\rho = uu^*, \quad . \quad . \quad . \quad . \quad (16)$$

continuously distributed throughout all spaces is to denote a density; when multiplied by m , it is to denote the *density of mass*; when multiplied by e the *density of charge*. (if our point-mass is an electron). This asserts that we get the correct electrodynamic actions (forces and emission of radiation) by calculating as if the charge of the electron is continuously distributed with the density ep in space. We shall discuss a possible physical meaning of this somewhat unattractive hypothesis later (§ 8). Here we shall straightway draw inferences from it.

~~If~~ ρ is a density, then

$$M = \int \rho p d\tau \quad . \quad . \quad . \quad . \quad (17)$$

is the "moment" of this distribution of density with respect to the co-ordinate q .

Thus M , when multiplied by e , denotes the *electric moment*, for example, for $q = x$, it denotes the x -component of the electric moment of our distribution of charge. Now we know that the amount and polarisation of the emitted radiation is determined electro-dynamically by calculating the time rate of change of the electric moment eM . Hence in the part of M that varies with the time we have a *measure of emission*.

But we must first take up a more general and more formal standpoint. For we must speak not only of *one* state and of the corresponding density ρ , but also of the transition from *one* state n to another m and of the "density" corresponding to this transition.

$$\rho_{nm} = u_n u_m^* \quad . \quad . \quad . \quad (18)$$

In a manner formally analogous with (17) we form the "moment" associated with this transition :

$$M_{nm} = \int q \rho_{nm} d\tau = \int q u_n u_m^* d\tau \quad . \quad . \quad . \quad (19)$$

In making these generalisations we, of course, lose sight of the original sense of the ideas "density" and "moment," so that these terms now serve only as guides to the analytical definitions contained in (18) and (19).

In (19) we substitute the expression of u given by (11), by applying it to two different proper functions n and m . We get

$$M_{nm} = q_{nm} e^{\frac{2\pi i}{h}(E_n - E_m)t} \quad . \quad . \quad . \quad (20)$$

$$q_{nm} = \int q \psi_n \psi_m^* d\tau \quad . \quad . \quad . \quad (21)$$

We call this quantity q_{nm} the "matrix element" of the co-ordinate q and so hint at its origin in the matrix-calculus. In the next section we shall show that the system of quantities q_{nm} introduced in (21) is identical with the q -matrix treated in § 4.

We first consider the *individual state* by setting $m = n$. Then, clearly, (18) becomes identical with (16) and (20) with (17). The density ρ and the moment M are in this case constant in time and the emission of radiation is nil. The *proper states* (the stationary orbits of the old theory) are *radiationless*.

We next consider the transition $n \rightarrow m$. The density ρ_{nm} and the moment M_{nm} are then variable in time. The vibration number of the moment and hence also of the emitted radiation is, by (20),

$$\nu = \frac{E_n - E_m}{h} \quad . \quad . \quad . \quad (22)$$

Here we have *Bohr's frequency condition*, which (cf. the beginning of this section), like the quantum condition, fits into wave-mechanics. We do

not pretend in this way to have *derived* the frequency condition, say on the basis of classical mechanics. For we had already included it essentially in the assumption (11) which expresses the dependence of the individual state on time. But it is important that we have now a scheme of calculation, moulded on the lines of the classical theory, which enables us to form judgments about the details of the radiation emitted, its polarisation and its intensity. For we conclude from (21) that if q_{nm} vanishes for certain members n and m and for a certain co-ordinate direction, for example, $q = x$, we have a *rule of polarisation*: the transition $n \rightarrow m$ yields no radiation corresponding to a vibration in the x -direction.

If for given values of n and m q_{nm} vanishes for every choice of $q (q = x, y, z)$, then the transition $n \rightarrow m$ is radiationless. We interpret this as meaning that this transition is *forbidden*. This gives us a *selection rule*.

If q_{nm} differs from zero, we regard $|q_{nm}|$ as a measure of the intensity for the transition in question and for the direction of polarisation. The ψ_n 's and ψ_m 's are, of course, to be imagined normalised in the sense of the general condition (15).

What holds for the transition $n \rightarrow m$ also holds for the transition $m \rightarrow n$. By (22) the latter transition is characterised by the quantity conjugate to q_{nm} :

$$q_{mn} = \int q \psi_m \psi_n^* d\tau = q_{nm}^* \quad . \quad . \quad . \quad (23)$$

so that

$$|q_{mn}| = |q_{nm}| \quad . \quad . \quad . \quad . \quad (24)$$

Our measure of intensity depends in a *symmetrical way* on the initial state n and the final state m —a very remarkable law which reflects itself in the summation rules of Burger and Dorgelo (*Atombau*, Chap. VIII, § 5).

The fact contained in eqn. (23) means in the language of § 4, eqn. (3a), that the matrix of the q 's is Hermitean.

In introducing the pseudo-moment M_{nm} and the matrix element q_{nm} we have left the original ground of wave-mechanics. Schrödinger, to whom we owe the representation given in (21) of the q -matrix, has attempted to account for their introduction by adapting them to his original ideas. He assumes that in the transition $n \rightarrow m$ both proper states n and m are to coexist in certain proportions which may be measured by the coefficients c_n and c_m . The state due to their superposition is then represented by

$$u = c_n \psi_n e^{\frac{2\pi i}{h} E_n t} + c_m \psi_m e^{\frac{2\pi i}{h} E_m t}$$

$$u^* = c_n^* \psi_n^* e^{-\frac{2\pi i}{h} E_n t} + c_m^* \psi_m^* e^{-\frac{2\pi i}{h} E_m t}$$

and we get from the original definition of density in (16) :

$$uu^* = c_n c_n^* \psi_n \psi_n^* + c_m c_m^* \psi_m \psi_m^* + c_n c_m^* \psi_n \psi_m^* e^{\frac{2\pi i}{h}(E_n - E_m)t} \\ + c_m c_n^* \psi_m \psi_n^* e^{\frac{2\pi i}{h}(E_m - E_n)t} \quad (25)$$

Thus this density is composed of parts that are constant in time—these do not interest us—and of two parts varying with time, that have the common vibration number (22).

Schrödinger then forms the moment of the density distribution (25) by the original rule (17) :

$$M = c_n c_n^* \int q \psi_n \psi_n^* d\tau + c_m c_m^* \int q \psi_m \psi_m^* d\tau + 2Rc_n c_m^* q_{nm} e^{\frac{2\pi i}{h}(E_n - E_m)t}. \quad (26)$$

Here R denotes "Real part of" (that is half the sum of the complex quantity which it precedes and its conjugate) and q_{nm} denotes the matrix element defined in (21); introduced in this way q_{nm} appears as a quantity which is characteristic for the state arising out of the superposition of the two proper vibrations. The further deductions regarding polarisation and intensity of emission are then the same as above.

We have chosen an abstract and formal representation of the rules of emission in preference to this apparently more picturesque* (*anschaulich*) introduction of q_{nm} for the following reason.

The assumption for u , which led to (25) and (26), contains the arbitrary coefficients c_n, c_m . These must reasonably be interpreted as excitation intensities (*Anregungsstärken*) of the states n and m . It is clear that the emitted intensity depends on the number of atoms in which the initial state n is excited, that is, on the coefficient c_n . But by eqn. (26) it would also depend on the coefficient c_m , that is, on the number of atoms in which the final state is realised. This seems meaningless and contradicts the results of experiments, for example, those in which the energy of electrons is transferred to atoms or molecules by collision (*Elektronenstoß*). It surely is not true that for $c_m = 0$, that is when the final state is not excited, the intensity of the transition $n \rightarrow m$ must vanish. Rather, there are "spontaneous" transitions also to non-excited states. We are therefore compelled to regard the necessity for introducing indefinite coefficients c_n, c_m , as a weak point in Schrödinger's representation of questions of intensity.

A really satisfactory treatment of these questions can be expected only when radiation has been fixed into the foundations of the new theory.† Following Schrödinger (and also Heisenberg) we have considered it sufficient to derive the radiation externally, as it were, from the moment

* Adopted from H. F. Biggs, *Wave Mechanics*, p. 15, footnote (Oxford University Press).

† Cf. an important paper by P. A. M. Dirac, *Proc. Roy. Soc.*, **114**, 248, 1927; in which, although by quite different methods, the spontaneous transitions have been successfully treated.

M or M_{nm} . Our only basis for this is the general correspondence with classical electrodynamics.

§ 6. Intensities in the case of Oscillators and Rotators

A. The Linear Harmonic Oscillator

We can now supplement our wave-mechanical treatment in § 3, A, so far that it becomes of equal value with the quantum-mechanical treatment in § 4.

Since in the case of the linear oscillator we have only one co-ordinate $q = x$, we write x_{nm} in place of q_{nm} , so that by eqn. (21) of the preceding section (the ψ 's are real)

$$x_{nm} = \int_{-\infty}^{\infty} x \psi_n \psi_m dx \quad . \quad . \quad . \quad (1)$$

Instead of x we use the more convenient dimensionless co-ordinate ξ from eqn. (5a) in § 3.

$$\xi = \sqrt{\alpha} x, \quad \alpha = \frac{2\pi m \omega_0}{h} \quad . \quad . \quad . \quad (2)$$

We then get

$$\alpha x_{nm} = \xi_{nm} = \int_{-\infty}^{+\infty} \xi \psi_n(\xi) \psi_m(\xi) d\xi \quad . \quad . \quad . \quad (3)$$

We show that ξ_{nm} differs from zero only if $m = n \pm 1$.

With this purpose in view we recall the analytical representation of the proper functions ψ_n . By eqn. (8) in § 3 we have

$$\psi_n = \frac{1}{N_n} H_n(\xi) e^{-\xi^2/2} \quad . \quad . \quad . \quad (4)$$

$H_n(\xi)$ denotes the Hermitean polynomial of the n^{th} degree, which was defined in § 3 by the recurrence formulæ of its coefficients. N_n is a normalising factor which is to be chosen so that ψ satisfies the general condition of normalisation (15) in § 5. Consequently [cf. (4) and (2)]

$$N_n^{-2} \int_{-\infty}^{+\infty} H_n^2(\xi) e^{-\xi^2} d\xi = \int_{-\infty}^{+\infty} \psi^2 d\xi = \sqrt{\alpha} \int_{-\infty}^{+\infty} \psi^2 dx = \sqrt{\alpha} \quad (5)$$

We next assume that

$$n < m$$

and write in place of (3)

$$N_n N_m \cdot \xi_{nm} = \int_{-\infty}^{+\infty} G_{n+1} H_m e^{-\xi^2} d\xi \quad . \quad . \quad (6)$$

in which we have set

$$G_{n+1}(\xi) = \xi H_n(\xi).$$

Thus G_{n+1} is a polynomial of the $(n+1)^{\text{th}}$ degree. But we can build up any polynomial of the $(n+1)^{\text{th}}$ degree out of the successive polynomials H in the form

$$G_{n+1}(\xi) = \sum_0^{n+1} c_r H_r(\xi) \quad . \quad . \quad . \quad (7)$$

§ 6. Intensities in the case of Oscillators and Rotators 49

just as well as out of the successive powers $\xi^0, \xi^1, \dots, \xi^{n+1}$. The coefficients c_ν are to be taken from it in such a way that all powers, from ξ^{n+1} to ξ^0 , are set equal to each other. Some of the coefficients c_ν can vanish (for example, in our case all c_ν 's with an even ν vanish if n is even). We insert the form (7) in (6) and get

$$N_n N_m \cdot \xi_{nm} = \sum_0^{n+1} c_\nu \int_{-\infty}^{+\infty} H_\nu H_m e^{-\xi^2} d\xi. \quad (8)$$

But from the condition of orthogonality (15) of the preceding section every term of this series vanishes, with the possible exception of the last one, namely in the case

$$m = n + 1 \quad (9)$$

In this case we have, on account of (5)

$$\xi_{n, n+1} = \frac{N_{n+1}}{N_n} c_{n+1} \sqrt{\alpha}. \quad (10)$$

The converse case

$$m < n$$

is dealt with by interchanging m and n . In place of (9) we thus find for the condition that ξ_{nm} may not vanish:

$$n = m + 1, \text{ that is, } m = n - 1, \quad (9a)$$

and in place of (10)

$$\xi_{n, n-1} = \frac{N_n}{N_{n-1}} c_n \sqrt{\alpha}. \quad (10a)$$

Only the trivial case

$$m = n$$

remains, in which

$$\xi_{nn} = \int_{-\infty}^{+\infty} \xi \psi_n^2 d\xi = 0 \quad (10b)$$

since ψ_n^2 is even, that is, $\xi \psi_n^2$ is odd in ξ .

This proves the *selection rule* for the harmonic oscillator. We find as in § 4, taking into account the reflections at the end of § 5: *all transitions $n \rightarrow m$ are forbidden with the exception of $m \rightarrow n \pm 1$.*

Passing on to questions of intensity we have to calculate the coefficients c_{n+1} and c_n by means of (10) and (10a). Let a_n be the coefficient of the highest power ξ^n in H_n . A comparison of the powers ξ^{n+1} on both sides of (7) at once shows that

$$a_n = c_{n+1} a_{n+1},$$

that is,

$$c_{n+1} = \frac{a_n}{a_{n+1}}, \quad c_n = \frac{a_{n-1}}{a_n}. \quad (11)$$

To calculate the a_n 's we require a formal representation of Hermit's polynomials.

We now assert (the proof is given just below) that

$$H_n(\xi) = (-1)^n e^{\xi^2} \frac{d^n e^{-\xi^2}}{d\xi^n} \quad (12)$$

This method of representation shows that the highest coefficient a_n , which is obtained by successive differentiation of $e^{-\xi^2}$ is

$$a_n = 2^n \quad (13)$$

It also shows that the integral

$$x = \int_{-\infty}^{+\infty} H_n(\xi) e^{-\xi^2} d\xi \quad (14)$$

which occurs in (5) has the value

$$x = 2^n n! \sqrt{\pi} \quad (15)$$

This is proved as follows: if we substitute for one of the two factors H in (14) the form (12), we get

$$x = (-1)^n \int \frac{d^n e^{-\xi^2}}{d\xi^n} H_n(\xi) d\xi,$$

and by integrating by parts n times, noticing that all derivatives of $e^{-\xi^2}$ vanish at the limits $\xi = \pm \infty$,

$$x = \int e^{-\xi^2} \frac{d^n H_n}{d\xi^n} d\xi.$$

If, as before, we insert

$$H_n \xi = a_n \xi^n + \dots$$

we get

$$x = a_n n! \int_{-\infty}^{+\infty} e^{-\xi^2} d\xi = a_n n! \sqrt{\pi},$$

and hence, on account of (13), $x = 2^n n! \sqrt{\pi}$, which was to be proved.

We have yet to add the proof that the polynomials H introduced in eqn. (12) satisfy eqn. (6) of § 3. Taking (7) of § 3 into consideration we may write this equation in the form

$$H_n'' - 2\xi H_n' + 2n H_n = 0 \quad (16)$$

The following simple proof may also be applied to other cases (spherical harmonics, Laguerre polynomials). We set

$$u = e^{-\xi^2} \quad (16a)$$

and form

$$\frac{du}{d\xi} = u' = -2\xi u.$$

By successive differentiation it follows from the formula for such differentiation in the case of a product that

$$u^{(n+2)} = -2\xi u^{(n+1)} - 2(n+1)u^{(n)} \quad (17)$$

§ 6. Intensities in the case of Oscillators and Rotators 51

But by (12)

$$(-1)^{n_u(n)} = e^{-\xi^2 H_n},$$

and so we get

$$\begin{aligned} (-1)^{n_u(n+1)} &= e^{-\xi^2(H'_n - 2\xi H_n)} \\ (-1)^{n_u(n+2)} &= e^{-\xi^2(H''_n - 4\xi H'_n + (4\xi^2 - 2)H_n)}. \end{aligned}$$

If we substitute these expressions in (17) we get (16), which we wished to prove.

The value of the normalising factor N_n now follows out of eqns. (14), (15), and (5), and is found to be

$$\frac{1}{N_n^2} 2^{nn}! \sqrt{\pi} = \sqrt{\alpha}, \quad N_n = \left(2^{nn}! \sqrt{\frac{\pi}{\alpha}}\right)^{-\frac{1}{2}}. \quad (18)$$

And so

$$\frac{N_n}{N_{n-1}} = \sqrt{2n}.$$

Further, it follows from (13) and (11) that

$$c_n = \frac{1}{2}.$$

By (10a) we therefore have

$$\xi_{n, n-1} = \sqrt{\frac{n\alpha}{2}}.$$

To return to our original measure of intensity $x_{n, m}$ we use eqn. (3) with $m = n - 1$ and take eqn. (2) into account. We find

$$x_{n, n-1} = \sqrt{\frac{n}{2\alpha}} = \sqrt{\frac{hn}{4\pi m\omega_0}}. \quad (19)$$

and, of course, simultaneously

$$x_{n, n+1} = x_{n+1, n} = \sqrt{\frac{h(n+1)}{4\pi m\omega_0}}. \quad (19a)$$

But these values are identical with the values $|q_{n, n-1}|$ and $|q_{n, n+1}|$ in eqn. (12), § 4. In the same way our eqn. (10b) which we may also write in the form

$$x_{n, n} = 0 \quad (19b)$$

is identical with eqn. (9a) of § 4, with $i = k = n$.

We thus demonstrate in our example what was generally established by Schrödinger* and Eckart,† namely that *the methods of quantum mechanics and wave-mechanics lead to the same results not only for the energy values and the proper values but also for the intensities.*

* E. Schrödinger, On the Relation between the Quantum Mechanics of Heisenberg, Born, and Jordan, and that of Schrödinger, pp. 45-61 of *Collected Papers*, Blackie & Son, Ltd. The original German paper is Ann. d. Phys., 79, 784, 1926.

† C. Eckart, Operator Calculus and the Solution of the Equations of Quantum Dynamics. Phys. Rev., 23, 711, 1926.

We now fill in a gap which was left at the end of § 4. It seemed there as if quantum mechanics could grasp the transitions but not the *states* of the oscillator. The states are first described in quantum mechanics by the diagonal elements $q_{n,n}$ of the matrix, which vanish in the case of the oscillator. In wave-mechanics these elements denote the "moments of the first order" of the charge distribution ψ^2 . But in wave-mechanics we may also form the "moments of higher order," for example

$$(x^2)_{n,n} = \alpha^{-3/2} \int_{-\infty}^{+\infty} \xi^2 \psi_n^2 d\xi \quad . \quad . \quad . \quad (19c)$$

This *quadratic* moment and all other moments of *even order* do *not* vanish. Correspondingly, the diagonal elements of the matrix of all even powers of q do not vanish. The totality of these moments can exactly serve to characterise the state in question in the language of quantum mechanics. [By a famous theorem due to Stieltjes any arbitrary distribution of masses $\psi(x)$ can be uniquely determined by the totality of their moments.] For example, we find for the integral in (19c) by the simple method given at the beginning of this section

$$N_n^2 (\xi^2)_{n,n} = \int G_{n+2} H_n e^{-\xi^2} d\xi = c_n,$$

where c_n denotes the coefficient of H_n in the expansion of $G_{n+2} = \xi^2 H_n$ in terms of the polynomials H_ν . This coefficient c_n may easily be calculated from the highest coefficients of the polynomials H_n , H_{n+1} , and H_{n+2} . It is obvious that this analytical method of calculating the quadratic moment is much simpler than the corresponding calculation of the matrix calculus. It is also clear that the description of a state by means of a single continuous function of state (phase function) ψ gives a clearer view than the description by means of the infinite series of moments to which quantum mechanics leads.

B. The Rotator in Space

Let the co-ordinates of the rotator be θ and ϕ as in § 3, B; the rectangular co-ordinates of the rotating point-mass—which we shall call ξ , η , ζ and at once connect them as complex variables—are then expressed as follows:

$$\xi + j\eta = a \sin \theta e^{j\phi}, \quad \zeta = a \cos \theta \quad . \quad . \quad . \quad (20)$$

We have written j for $\pm i$ in order to distinguish the indefiniteness of sign that occurs here from one that occurs later. The proper functions are

$$\psi_{m,\mu} = \frac{1}{N_\theta} \frac{1}{N_\phi} P_m^\mu(\cos \theta) e^{\pm i\mu\phi} \quad . \quad . \quad . \quad (21)$$

The "associated spherical harmonics" P_m^μ are taken as normalised in the usual way, namely in accordance with the eqns. (12) and (12a) of § 2. On the other hand we take our ψ as normalised in a rational way "to unity," according to eqn. (15b), § 5, on account of which we added the normalising factors N_θ , N_ϕ in (21). We determine N_θ and N_ϕ in such a

§ 6. Intensities in the case of Oscillators and Rotators 53

way that we normalise the P_m^μ 's and the $e^{i\mu\phi}$'s separately to unity, that is, we not only demand that

$$\int \psi_{m,\mu} \psi_{m,\mu}^* d\sigma = 1, \quad d\sigma = \sin \theta \, d\theta \, d\phi \quad (22)$$

but also that

$$\int_{-1}^{+1} [P_m^\mu(x)]^2 dx = N_\theta^2, \quad x = \cos \theta \quad (23)$$

and

$$\int_0^{2\pi} e^{i\mu\phi} e^{-i\mu\phi} d\phi = N_\phi^2 = 2\pi \quad (23a)$$

The conditions of orthogonality require [cf. (15a) in § 5, where now the simple indices n and m are to be replaced by the double indices m, μ and m', μ'] that

$$\int \psi_{m,\mu} \psi_{m',\mu'}^* d\sigma = 0,$$

except when $\mu = \mu'$ and at the same time $m = m'$. For $\mu \neq \mu'$ this equation, in virtue of the integral over ϕ , is fulfilled; for $\mu = \mu'$ it requires that

$$\int_{-1}^{+1} P_m^\mu(x) P_{m'}^\mu(x) dx = 0, \quad m \neq m' \quad (24)$$

It must be particularly emphasised that we do not need to prove this equation by means of special calculations with spherical harmonics, but that it follows immediately from the general theorem of the previous section on the orthogonality of proper functions.

To calculate the normalising factor N_θ we must have a formula to express the P_m^μ 's. The following is best suited to our purpose:

$$P_m^0(x) = P_m(x) = \frac{1}{2^m m!} \frac{d^m}{dx^m} (x^2 - 1)^m \quad (25)$$

from which, by (12) of § 2, it follows that

$$P_m^\mu(x) = \frac{(1 - x^2)^{\mu/2}}{2^m m!} \frac{d^{m+\mu}}{dx^{m+\mu}} (x^2 - 1)^m \quad (25a)$$

It is immediately clear that (25) fulfils the normalising condition (12a) in § 2, but we must also show that (25) satisfies the differential eqn. (2) of § 2, which, with $y = P_m$, runs in our present symbols:

$$(1 - x^2)y'' - 2xy' + m(m+1)y = 0 \quad (26b)$$

Similarly as in (16a), we set $u = (x^2 - 1)^m$. By a single differentiation we get

$$(x^2 - 1)u' = 2mxu$$

and by successive differentiation $m + 1$ times

$$(x^2 - 1)u^{(m+2)} + 2(m+1)xu^{(m+1)} + m(m+1)u^{(m)} \\ = 2mxu^{(m+1)} + 2m(m+1)u^{(m)}.$$

If we collect terms and write $u^{(m)} = y$, by (25), then (26b) actually results.

To calculate the integral that occurs in (23), that is, the square of the normalising factor N_θ , we form, by (25a)

$$N_\theta^2 = \frac{1}{2^{2m}(m!)^2} \int_{-1}^{+1} G_{m+\mu} \frac{d^{m+\mu}}{dx^{m+\mu}} (x^2 - 1)^m dx \quad (27)$$

in which

$$G_{m+\mu} = (1 - x^2)^\mu \frac{d^{m+\mu}}{dx^{m+\mu}} (x^2 - 1)^m;$$

the highest term in G is $ax^{m+\mu}$ with

$$a = (-1)^\mu 2m(2m-1) \dots (m-\mu+1) = (-1)^\mu \frac{(2m)!}{(m-\mu)!} \quad (28)$$

By successive integration by parts we get from (27)

$$N_\theta^2 = (-1)^{m+\mu} \frac{(m+\mu)!}{2^{2m}(m!)^2} \int_{-1}^{+1} (x^2 - 1)^m dx \quad (27a)$$

[The factors $1 \pm x$ that occur just suffice to make the terms without integral signs vanish.]

Now

$$q_m = \int_{-1}^{+1} (x^2 - 1)^m dx = \int_{-1}^{+1} (x^2 - 1)^{m-1} x^2 dx - \int_{-1}^{+1} (x^2 - 1)^{m-1} dx \\ = \frac{1}{2m} \int_{-1}^{+1} x \frac{d}{dx} (x^2 - 1)^m dx - q_{m-1} = -\frac{1}{2m} q_m - q_{m-1}$$

satisfies the following recurrence formula:

$$q_m = -\frac{2m}{2m+1} q_{m-1}.$$

On account of $q_0 = 2$ it follows from this formula that

$$q_m = (-1)^m \frac{2m(2m-2)(2m-4) \dots 2}{(2m+1)(2m-1)(2m-3) \dots 3} \cdot 2 = (-1)^m 2 \cdot \frac{(2^m m!)^2}{(2m+1)!}$$

and hence from (27a) and (28) that

$$N_\theta^2 = \frac{2}{2m+1} \frac{(m+\mu)!}{(m-\mu)!} \quad (30)$$

and, in particular, for $\mu = 0$ we get the well-known formula

$$N_\theta^2 = \frac{2}{2m+1} \quad (30a)$$

§ 6. Intensities in the case of Oscillators and Rotators 55

The total normalising factor, which is equal to the product of N_ϕ and N_θ and is to be denoted by $N_{m, \mu}$ is given, by eqns. (30) and (23a), by

$$N_{m, \mu}^2 = \frac{4\pi}{2m+1} \frac{(m+\mu)!}{(m-\mu)!} \quad (31)$$

We now come to the question of intensity and polarisation in the case of the rotator. The basis is again given by the matrix elements of the co-ordinates, namely of the rectangular co-ordinates (20). Logically, we should now write these with four elements:

$$\zeta_{m\mu, m'\mu'} \quad \text{and} \quad (\xi + i\eta)_{m\mu, m'\mu'}.$$

But we easily show that the former and the latter differ from zero only if

$$\mu' = \mu \quad \text{or} \quad \mu' = \mu \pm 1, \text{ respectively} \quad (32)$$

so that after the respective value of μ' has been inserted we may adopt the abbreviated notation

$$\zeta_{mm'} \dots (\mu' = \mu) \quad \text{and} \quad (\xi + j\eta)_{mm'} \dots (\mu' = \mu \pm 1), \quad (32a)$$

in which the two values $\mu' = \mu \pm 1$ may be associated with both signs of j .

The proof of (32) consists in the fact that the matrix elements for ζ or $\xi + j\eta$ contain the following integrals of ϕ respectively:

$$\int_0^{2\pi} e^{\pm i(\mu - \mu')\phi} d\phi \quad \text{or} \quad \int_0^{2\pi} e^{j\phi} e^{\pm i(\mu - \mu')\phi} d\phi.$$

They do *not* vanish only if

$$\mu - \mu' = 0 \quad \text{or} \quad \mu - \mu' = \pm 1$$

respectively, and then their value is 2π . This proves the selection rule (32).

But this has simultaneously furnished us with a "rule of polarisation" in the following way. The case $\mu' = \mu$ corresponds to a vibration parallel to the ζ -axis (a variable electric moment M in this direction, cf. § 5): the case $\mu' = \mu \pm 1$ denotes a right-handed or left-handed circularly polarised vibration in the ξ, η -plane (corresponding to an electric moment in this plane variable in direction but not in magnitude). Moreover, we must remark that the ζ -axis can be empirically distinguished from the other directions only if it is physically distinguished, for example by a magnetic field. Our rule of polarisation would accordingly come into action only in the case of the Zeeman effect, which, however, for the rotator consisting of the diatomic molecule, as realised in band spectra, becomes so small that so far it has not been observed. For the rest, our present rule of polarisation is identical with that which applies to the normal Zeeman effect of the electron (I, Chap. V, § 6, p. 294). Cf. also the treatment of the Zeeman effect by Wave-Mechanics in § 10.

We come to the selection rule for the m 's and show that it runs

$$m' = m \pm 1 \quad (33)$$

To do this we must actually work out the quantities $\zeta_{m,m'}$ and $(\xi + j\eta)_{m,m'}$.

Replacing the integrals over ϕ by 2π , we obtain by (20) and (21)

$$N_{m\mu} N_{m'\mu} \zeta_{mm'} = 2\pi a J, \quad N_{m\mu} N_{m'\mu} (\xi + j\eta)_{mm'} = 2\pi a K \quad (33a)$$

$$\left. \begin{aligned} J &= \int_0^\pi \cos \theta P_m^\mu(\cos \theta) P_{m'}^\mu(\cos \theta) \sin \theta d\theta \\ K &= \int_0^\pi \sin \theta P_m^\mu(\cos \theta) P_{m'}^{\mu \pm 1}(\cos \theta) \sin \theta d\theta \end{aligned} \right\} \quad (33b)$$

We insert the variable $x = \cos \theta$ in J and K , and note that by (25a) the P_m^μ 's are equal to $(1 - x^2)^{\mu/2}$ multiplied by a polynomial of the $(m - \mu)$ th degree in x , which we shall call $P_{m-\mu}^\mu$;* further, we introduce two polynomials F and G , of degree indicated by the lower index. Thus we set

$$\left. \begin{aligned} P_m^\mu(\cos \theta) &= (1 - x^2)^{\mu/2} P_{m-\mu}^\mu(x), \\ \sin \theta P_{m'}^{\mu \pm 1}(\cos \theta) &= (1 - x^2)^{\mu/2} (1 - x^2)^{1/2} \pm 1/2 P_{m'-\mu \mp 1}^{\mu \pm 1}(x), \\ F_{m'-\mu+1} &= x P_{m-\mu}^\mu, \quad G_{m'-\mu+1} = (1 - x^2)^{1/2} \pm 1/2 P_{m-\mu}^{\mu \pm 1} \end{aligned} \right\} \quad (34)$$

This leads to the expressions

$$J = \int_{-1}^{+1} F \cdot P_{m-\mu}^\mu (1 - x^2)^\mu dx, \quad K = \int_{-1}^{+1} G \cdot P_{m-\mu}^\mu (1 - x^2)^\mu dx \quad (35)$$

The somewhat artificial definition of G in (34) is due to the fact that in the integral K we had to extract the factor $1 - x^2$ raised to the power indicated by the upper index of the preceding P .

We can now use exactly the same line of argument as under A in eqn. (6). Build up F and G from the polynomials P , instead of from the successive powers of x , thus:

$$F = \sum_0^{m'-\mu+1} c_\nu P_\nu^\mu, \quad G = \sum_0^{m'-\mu+1} d_\nu P_\nu^\mu \quad (35a)$$

and assume

$$m > m'.$$

By substituting (35a) in (35) the condition of orthogonality (24) shows us that $J = K = 0$ unless

$$m' = m - 1.$$

This proves one part of our selection rule (33). At the same time we have, on account of the normalising condition (23),

$$J = \frac{c_{m-\mu}}{2\pi} N_{m,\mu}^2, \quad K = \frac{d_{m-\mu}}{2\pi} N_{m,\mu}^2.$$

* Similar to the notation used by E. Heine, *Handbuch der Kugelfunktionen*, Chap. IV, § 46.

§ 6. Intensities in the case of Oscillators and Rotators 57

and hence by (33a)

$$\zeta_{m, m-1} = a \frac{N_{m, \mu}}{N_{m-1, \mu}} c_{m-\mu}, \quad (\xi + j\eta)_{m, m-1} = a \frac{N_{m, \mu}}{N_{m-1, \mu \pm 1}} d_{m-\mu} \quad (36)$$

From (34) and (35a), however, if $p_{m, \mu}$ denotes the coefficient of the highest power in $P_{m-\mu}^{\mu}$, namely in consequence of (25a):

$$\left. \begin{aligned} p_{m, \mu} &= \frac{(2m)!}{2^m m! (m-\mu)!} \\ c_{m-\mu} &= \frac{p_{m-1, \mu}}{p_{m, \mu}} = \frac{m-\mu}{2m-1} \\ d_{m-\mu} &= \mp \frac{p_{m-1, \mu \pm 1}}{p_{m, \mu}} = \frac{-(m-\mu-1)(m-\mu)}{2m-1} \cdot \frac{1}{2m-1} \end{aligned} \right\} \quad (36a)$$

If we substitute from (36a) and (31) in (36), calculation gives

$$\left. \begin{aligned} \zeta_{m, m-1} &= a \sqrt{\frac{(m+\mu)(m-\mu)}{(2m+1)(2m-1)}} \dots (\mu' = \mu) \\ (\xi + j\eta)_{m, m-1} &= -a \sqrt{\frac{(m-\mu-1)(m-\mu)}{(2m+1)(2m-1)}} \dots (\mu' = \mu + 1), \\ (\xi + j\eta)_{m, m-1} &= a \sqrt{\frac{(m+\mu-1)(m+\mu)}{(2m+1)(2m-1)}} \dots (\mu' = \mu - 1). \end{aligned} \right\} \quad (37)$$

With the converse assumption

$$m < m'$$

we find for the condition that J and K should not vanish by interchanging m and m'

$$m' = m + 1$$

as the second part of our selection rule (33). The expressions (37) then become changed in that $m, m+1$ now take the place of $m-1, m$.

To make the construction of the formulæ (37) intelligible we shall divide the transitions in question into "parallel," "anti-parallel," and "neutral" transitions (*gleichsinnig, gegensinnig und neutral*), analogously to *Atombau*, p. 580. A transition is to be called *parallel* when both numbers m and μ change in the same sense, for example, $m \rightarrow m-1$ and $\mu \rightarrow \mu-1$; they are to be called *anti-parallel* when their values move in opposite directions, for example, $m \rightarrow m-1$ and $\mu \rightarrow \mu+1$; a *neutral* transition occurs when μ remains unaltered. The transitions $m \rightarrow m-1$ in (37) are then respectively: neutral, anti-parallel and parallel. Just as in I, Chap. VI, p. 367, the parallel transitions are "strong," the anti-parallel "weak," and the neutral "medium" (less strong). In our formulæ (37) this is expressed in the fact that in the strong transition (third row) $+\mu, +\mu$, in the weak transition (second row) $-\mu, -\mu$ and in the medium transition (first row), $+\mu, -\mu$ occur under the root sign.

Since the individual μ -components, on account of their insufficient resolution, do not interest us, we shall sum up over all the μ 's, taking the sum for the *intensities* and not for the amplitudes, with due regard to the fact that the proper states distinguished by the values of μ must be considered as *incoherent* among themselves. Thus we sum up the *squares* of ζ and of the *absolute values* (*Normen*) of $\xi + j\eta$. In performing this summation it is convenient to let μ go from $-m$ to $+m$,* and then, strictly speaking, $|\mu|$ should be written for μ in the upper index of the spherical harmonics and in the formulæ (37).

By the formula

$$\sum_{-m}^{+m} \mu^2 = \frac{m}{3}(m+1)(2m+1) \quad . \quad . \quad (38)$$

we get from (37) that

$$\sum \zeta^2 = a^2 \sum_{-m}^{+m} \frac{m^2 - \mu^2}{(2m+1)(2m-1)} = \frac{a^2}{3} m \quad . \quad . \quad (38a)$$

and, taking together the two formulæ (37) for $\xi + j\eta$:

$$\sum |\xi + j\eta|^2 = 2a^2 \sum_{-m}^{+m} \frac{m(m-1) + \mu^2}{(2m+1)(2m-1)} = \frac{4a^2}{3} m \quad . \quad (38b)$$

The sum † of (38a) and (38b) gives

$$\frac{1}{a^2} J = m \quad . \quad . \quad . \quad (39)$$

J is the total intensity which is observed for the transition $m \rightarrow m-1$, that is, of a definite line in the spectrum of our rotator. In the same way we should obtain for the transition $m+1 \rightarrow m$, which gives rise to another line of the spectrum:

$$\frac{1}{a^2} J = m+1 \quad . \quad . \quad . \quad (39a)$$

The sum of (39) and (39a) gives $2m+1$. $2m+1$ is the "weight" of the state m (*Atombau*, Chap. VIII, § 5, p. 651), that is, the number of different states ($\mu = -m$ to $\mu = +m$) that belong to the same m , or we may also express the same thing by saying the number of spherical harmonics (cf. § 2, p. 13) that have the same lower index m . The

* Calculating with negative μ 's means nothing more than that $e^{-i\mu\phi}$ is to be considered equally with $e^{+i\mu\phi}$. Later on, the simplification effected by counting up the negative μ 's as well will again be found useful.

† The reason for counting only one half of (38b) and not the full amount, as might suggest itself by (37) on account of the two possibilities of transition $\mu+1 \rightarrow \mu$ and $\mu-1 \rightarrow \mu$, is that of the two circular vibrations in the $\xi\eta$ -plane, in each case only half the intensity is observed if, for example, the line of vision lies in the $\xi\eta$ -plane.

appearance of the weight $2m + 1$ in the sum of the intensities of all transitions is in agreement with the summation rules of Burger and Dorgelo (*Atombau*, Chap. VIII, § 5). Our formulæ (39) and (39a) have been derived by Fowler* for the somewhat more general case of the rotation-vibration spectrum by means of the correspondence principle, and have been extended by Hönl and London† to the band-spectra with zero branches, whereas our treatment is restricted to the case of the pure rotation spectrum. We wish to point out again, however, that our treatment on the basis of wave-mechanics renders superfluous the introduction of the classical theory, by way of the correspondence principle, even in questions relating to intensity.

The oscillating rotator would naturally follow on the rotator at this point. We investigated its proper values in § 3, D. Its spectrum forms the type of the rotation vibration spectra just mentioned. But the investigation of its intensity conditions would diverge only little from those appertaining to the rigid oscillator, so that we may omit the discussion here and refer to the detailed investigation of Fues.‡

§ 7. The Kepler Problem

We now come to the central problem of wave-mechanics—the problem of the hydrogen atom. This formed the test of the method of wave-mechanics in the first paper on the subject by Schrödinger, published in 1926. In comparing it with the working out of the hydrogen atom by quantum-mechanical methods, carried out by Pauli§ at the same time, the superiority of wave-mechanics as regards mathematical simplicity and clearness becomes manifest. We hope in the sequel to enhance the lucidity of treatment still further.

A. Proper Values and Proper Functions. Discrete and Continuous Spectrum

The potential energy between an electron and a nucleus carrying a charge Ze is, if normalised in the usual way ($V = 0$ for $r = \infty$),

$$V = -\frac{Ze^2}{r}.$$

The wave equation (11) of § 1 becomes, if E is similarly normalised,

$$\Delta\psi + \frac{8\pi^2m}{h^2}\left(E + \frac{Ze^2}{r}\right)\psi = 0 \quad . \quad . \quad . \quad (1)$$

We treat the equation in *spatial* polar co-ordinates r, θ, ϕ and assume a solution in the form

$$\psi = R P_l^m(\cos \theta) e^{im\phi} \quad . \quad . \quad . \quad (2)$$

* R. H. Fowler, *Phil. Mag.*, **49**, 1272 (1925).

† H. Hönl and F. London, *Zeitschr. f. Phys.*, **33**, 803 (1925).

‡ E. Fues, *Ann. d. Phys.*, **81**, 281 (1926).

§ W. Pauli, jr., *Zeitschr. f. Phys.*, **36**, 336 (1925).

R is a function of r alone. As the lower index of the spherical harmonic P we have chosen l , for reasons connected with the mapping out of spectra (instead of the n in § 2 and the m in § 6). l must be an integer ≥ 0 and m must likewise be an integer, in order that (2) may represent a proper function (cf. § 2). If we use the differential equation of spherical harmonics in the form (1 b), § 2 with the proper value λ of eqn. (11), § 2, and derive the expression $\Delta\psi$ from eqn. (1) § 2, then we get for R the differential equation :

$$\frac{d^2R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left(A + 2\frac{B}{r} + \frac{C}{r^2} \right) R = 0 \quad (3)$$

where, for brevity, we have written

$$A = \frac{8\pi^2m}{h^2}E, \quad B = \frac{4\pi^2m}{h^2}Ze^2, \quad C = -l(l+1) \quad (3a)$$

We distinguish between the two cases

$$E < 0 \text{ and } E > 0.$$

(a) $E < 0$, *Corresponding to the Elliptic Orbits of the Earlier Theory*

To take account simultaneously of the sign and dimensions of A we set

$$A = -\frac{1}{r_0^2} \quad (4)$$

and determine the asymptotic behaviour of R by omitting all terms in $\frac{1}{r}$ and $\frac{1}{r^2}$ in (3). We obtain

$$\frac{d^2R}{dr^2} = \frac{R}{r_0^2}, \quad R_1 = e^{\pm r/r_0}.$$

Of these two asymptotic solutions we can use only that which vanishes at infinity. We introduce the dimensionless quantity

$$\rho = 2\frac{r}{r_0} = 2\sqrt{-A} \cdot r, \quad 0 \leq \rho \leq \infty \quad (4a)$$

that is, we write our asymptotic solution as

$$R = e^{-\rho/2}.$$

Corresponding to this, we make the general assumption

$$R = e^{-\rho/2} \cdot v \quad (5)$$

Multiplying eqn. (3) by $r^2/4$ and expressing it in terms of ρ (dashes denote differentiations with respect to ρ)

$$R'' + \frac{2}{\rho}R' + \left(-\frac{1}{4} + \frac{B}{\sqrt{-A}\rho} - \frac{l(l+1)}{\rho^2} \right) R = 0 \quad (6)$$

From (5) we calculate

$$R' = e^{-\rho/2}(v' - \frac{1}{2}v), \quad R'' = e^{-\rho/2}(v'' - v' + \frac{1}{2}v).$$

Substituting in (6), we get the differential equation for v :

$$v'' + \left(\frac{2}{\rho} - 1\right)v' + \left[\left(\frac{B}{\sqrt{-A}} - 1\right)\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right]v = 0 \quad (7)$$

The only singular point of this equation situated in finite regions is the limiting point $\rho = 0$: by criterion (5) of § 2 it is a pole. We therefore make the assumption

$$v = \rho^\gamma w, \quad w = \sum_{\nu=0,1,2,\dots} a_\nu \rho^\nu \quad (7a)$$

To find the characteristic exponent γ , we substitute (7a) in (7) and look for the factor of $a_0 \rho^{\gamma-2}$. It is

$$\gamma(\gamma-1) + 2\gamma - l(l+1) = \gamma(\gamma+1) - l(l+1) \quad (7b)$$

Equated to zero, it gives the two values $\gamma = l$ and $\gamma = -l-1$, of which, however, we can use only the first if our solution is to be a proper function of the problem. So we replace (7a) by

$$v = \rho^l w = \sum_{\nu=0,1,2,\dots} a_\nu \rho^{\nu+l} \quad (7c)$$

On substitution in (7) the factors of all powers of ρ must vanish. This gives a recurrence formula for the coefficients a_ν , namely, one consisting of two members. For we get from the factor of $\rho^{\nu+l-1}$

$$\left\{ \begin{aligned} & \{(\nu+l+1)(\nu+l) + 2(\nu+l+1) - l(l+1)\}a_{\nu+1} \\ & = \left\{ \nu+l+1 - \frac{B}{\sqrt{-A}} \right\} a_\nu \end{aligned} \right\} \quad (8)$$

We now wish to arrange that w becomes a polynomial, that is that the series (7c) ceases, say, at

$$\nu = n_r \quad (9)$$

in which the notation n_r may suggest "radial quantum number." We accomplish this by making the factor of a_ν in (8) zero for $\nu = n_r$ by setting

$$\frac{B}{\sqrt{-A}} = n_r + l + 1 = n \quad (9a)$$

whereupon obviously all the later coefficients a_ν vanish for $\nu > n_r$. The symbol n is to hint at the "principal quantum number." From (9a) we now get, by squaring,

$$-\frac{B^2}{A} = n^2.$$

But in view of the meaning of A and B in (3a) this is the *Balmer term* and includes the correct definition of the *Rydberg constant* $2\pi^2 m e^4 / h^3$:

$$E = - \frac{2\pi^2 m e^4 Z^2}{h^2 n^2} \quad (10)$$

Thus the *discrete spectrum* of the hydrogen problem has been found.

We have next to make some observations about the significance of our definition of l . The old principal quantum number n was composed, according to eqn. (9a), of two positive integers n_r and l , both of which could assume all values between zero and infinity; n_r , as already mentioned, corresponds to the radial quantum, l is our present "azimuthal quantum," such that $l + 1$ equals our earlier number k . Thus for k wave-mechanics gives all the positive integers *except zero*. If we interpret k in the sense of the old ideas of orbits, this means that the "pendulum orbits" (*Pendelbahnen*) $k = 0$, in executing which the electron would come infinitely near the nucleus, are excluded in the new theory quite automatically as being incompatible with the "boundary conditions." For the old quantum theory the reason for the exclusion of $k = 0$ offered an essentially insoluble problem,* for the exclusion of pendulum orbits in the case of the relativistic Kepler problem and of the Stark effect led to difficulties in the case of the Zeeman effect, according to the adiabatic principle (cf. *Atombau*, Chap. V, p. 405), where the corresponding orbits cannot be logically excluded. Particular disagreement was found to occur in the case of crossed electric and magnetic fields. Wave-mechanics overcomes these difficulties in the simplest fashion by not yielding those states which would correspond to the pendulum orbits. They do not come into consideration at all and so cannot conflict with the adiabatic principle. We may also point out that an appropriate representation of the facts offered by the anomalous Zeeman effect compelled us even in *Atombau*, Chap. VIII, to adopt measures which correspond exactly to the introduction of $l = k - 1$ which has its logical foundation in wave-mechanics.†

Further, we see that in the sense of the definition on p. 18, the Kepler problem is degenerate, for to the proper function with the quantum number n there correspond all proper functions

$$\psi = R_{n_r}(\rho_n) P_l^m(\cos \theta) e^{im\phi},$$

where for the present we denote the degree of R by the lower index n_r . Now the quantum numbers n , l , m , are bound together by the following relationships‡:

$$l \leq n - 1, \quad |m| \leq l,$$

as may be seen directly from the definition of P_l^m and from eqn. (9a); by

* See in particular the remarks by W. Pauli in Vol. XXIII of the *Handbuch der Physik* (Geiger-Scheel), especially pp. 159 *et seq.* and 164.

† For in *Atombau*, p. 587, eqn. (5), we defined $j_a = k - 1$, and saw in eqn. (18), p. 692, that the Landé g -formula, if the symmetry of its structure is not to be destroyed, demands the use of the j_a in place of k . We shall of course in future drop the rather unfortunately chosen symbol j_a and write l instead.

‡ Concerning the possibility of m having the negative sign see footnote, p. 58. In the sequel we shall as a rule calculate on the assumption that m takes all values from $-l$ to $+l$.

(9a) n_r is determined by l , if n is known. So for one and the same n we have the following number of proper functions :

$$\sum_{l=0}^{n-1} \sum_{m=-l}^{+l} m = \sum_{l=0}^{n-1} (2l+1) = n^2. \quad (10a)$$

that is, every proper value is $(n^2 - 1)$ -fold degenerate.

We wish to study the proper functions belonging to the proper values (10) more closely. By (5) and (7c) their radial component is

$$R = \rho^l \cdot w \cdot e^{-\rho/2},$$

where w is a polynomial of degree n_r . Although this polynomial is completely determined by the recurrence formula (8) (except for a multiplying constant), we shall set up the differential equation for w . For this purpose we calculate from (7c)

$$v' = \rho^l \left(w' + \frac{l}{\rho} w \right), \quad v'' = \rho^l \left(w'' + \frac{2lw'}{\rho} + \frac{l(l-1)}{\rho^2} w \right),$$

and by substituting in (7) and reducing simply with the help of (9a) we get

$$\rho w'' + [2(l+1) - \rho] w' + (n-l-1)w = 0. \quad (11)$$

We have already encountered this equation in § 3, eqn. (21b). As worked out there, it arises through successive differentiation from the differential equation of the " k th Laguerre polynomial"

$$xy'' + (1-x)y' + ky = 0 \quad (12)$$

For if we denote the i th differential quotient of y by w , then

$$xw'' + (i+1-x)w' + (k-i)w = 0. \quad (12a)$$

holds. Comparison with (11) shows that in our case

$$i = 2l+1, \quad k = n+l, \quad x = \rho.$$

Denoting the Laguerre polynomial as usual by L , its degree by the lower index and the number of successive differentiations by the upper index, we write

$$w = L_{n+l}^{(2l+1)}(\rho), \quad R = \rho^l \cdot L_{n+l}^{(2l+1)} \cdot e^{-\rho/2} / N_r. \quad (13)$$

N_r stands for a normalising factor that we shall presently calculate. Accordingly, the degree of w is

$$n+l-(2l+1) = n-l-1 = n_r$$

in agreement with (9) and (9a).

Next we shall derive a method of representing the Laguerre polynomials which is convenient for our purpose. We assert that if we again denote the degree by k

$$L_k(x) = e^x \frac{d^k}{dx^k} (x^k e^{-x}) \quad (14)$$

To prove this we set $u = x^k e^{-x}$

By differentiating once with respect to x we get

$$xu' = (k-x)u,$$

and from this by differentiating $(k+1)$ times and applying the product rule

$$xu^{(k+2)} + (k+1)u^{(k+1)} = (k-x)u^{(k+1)} - (k+1)u^{(k)}. \quad (14a)$$

If we now set $L_k = y$ and, linking up with (14), $u^{(k)} = ye^{-x}$ and work out

$$u^{(k+1)} = (y' - y)e^{-x}, \quad u^{(k+2)} = (y'' - 2y' + y)e^{-x}$$

we get by substitution in (14a)

$$x(y'' - 2y' + y) + (k+1)(y' - y) = (k-x)(y' - y) - (k+1)y$$

and this is in fact identical with (12).

By (13) the corresponding formula for our polynomial w is

$$w = \frac{d^{2l+1}}{d\rho^{2l+1}} \left(e^{\rho} \frac{d^{n+l}}{d\rho^{n+l}} (\rho^n + l e^{-\rho}) \right) \quad (15)$$

We now write down the condition of orthogonality for any two proper functions ψ_{nlm} and $\psi_{n'l'm'}$ according to eqn. (15a), § 5:

$$\int \psi_{nlm} \psi_{n'l'm'}^* d\tau = 0 \quad \left\{ \begin{array}{l} \text{except when } n = n', l = l' \text{ and } m = m' \\ \text{simultaneously.} \end{array} \right.$$

Since by § 6, B this condition is already fulfilled by the ϕ - and θ -components of the ψ 's, unless $m = m'$ and $l = l'$ simultaneously, in order to find the additional condition for the r -component we must set $m' = m$ and $l' = l$. We get (by definition, R is real) as the *radial condition of orthogonality*

$$\int R_{nlm} R_{n'l'm} r^2 dr = 0, \quad n' \neq n \quad (16)$$

Finally, for $n' = n$, $l' = l$ and $m' = m$ we get from eqn. (15b) of § 5 a *normalising condition for the functions R* . Taking into account the method of representing R in (13) and the changed notation for the quantum number, we get for it

$$\left(\frac{r_0}{2} \right)^3 J = N_r^2 \quad (17)$$

$$J = \int_0^\infty \rho^{2(l+1)} e^{-\rho} [L_{n+l}^{(2l+1)}(\rho)]^2 d\rho \quad (17a)$$

(17) determines the normalising factor N_r ; the factor $\left(\frac{r_0}{2} \right)^3$, which arises through the transition (4a) from r to ρ is due to the circumstance that our normalising condition refers to the variable r while we have obviously to effect our integration in the variable ρ .

To calculate J we express *one* of the two factors L as a differential quotient and combine *the other* with the power of ρ into a whole function G .

$$J = \int_0^\infty G_{n+l+1} \cdot e^{-\rho} \cdot \frac{d^{2l+1} L_{n+l}^{(2l+1)}}{d\rho^{2l+1}} d\rho$$

$$G_{n+l+1} = \rho^{2(l+1)} L_{n+l}^{(2l+1)} = a\rho^{n+l+1} + b\rho^{n+l} + \dots \quad (17b)$$

Integrating $(2l + 1)$ -times by parts and again writing down only the highest powers of ρ we obtain (the factors $e^{-\rho}$ and $\rho^{2(l+1)}$ make the terms not under the integral sign vanish)

$$J = \int_0^\infty \{a\rho^{n+l+1} + b\rho^{n+l} - a(2l+1)(n+l+1)\rho^{n+l} + \dots\} e^{-\rho} L_{n+l} d\rho,$$

that is, if we use the form (14) for L_{n+l}

$$J = \int_0^\infty \{a\rho^{n+l+1} + [b - a(2l+1)(n+l+1)]\rho^{n+l} + \dots\} \frac{d^{n+l}}{d\rho^{n+l}} (\rho^{n+l} e^{-\rho}) d\rho.$$

Integrating by parts $(n+l)$ -times more we get

$$\left. \begin{aligned} (-1)^{n+l} J &= a(n+l+1)! \int_0^\infty \rho^{n+l+1} e^{-\rho} d\rho \\ &+ \{b - a(2l+1)(n+l+1)\} (n+l)! \int_0^\infty \rho^{n+l} e^{-\rho} d\rho \end{aligned} \right\} \quad (17c)$$

while all the later terms of the expansion vanish in the integration by parts. (17c) is equivalent to

$$(-1)^{n+l} J = a[(n+l+1)!]^2 + \{b - a(2l+1)(n+l+1)\} [(n+l)!]^2 \quad (17d)$$

By (17b) and (14) the values of a and b are

$$a = (-1)^{n+l} \frac{(n+l)!}{(n-l-1)!}, \quad b = (-1)^{n+l-1} (n+l) \frac{(n+l)!}{(n-l-2)!}$$

Substituting in (17d) we get

$$J = \frac{2n[(n+l)!]^2}{(n-l-1)!} \quad (18)$$

This determines our normalising factor N_r . For by (17)

$$\frac{1}{N_r^2} = \left(\frac{2}{r_0}\right)^2 \frac{(n-l-1)!}{2n[(n+l)!]^3} \quad (18a)$$

(b) $E > 0$, Corresponding to the Hyperbolic Orbits of the Earlier Theory

We now set

$$A = + \frac{1}{r_0^2} \quad (19)$$

and get asymptotically by (3)

$$\frac{d^2 R}{dr^2} + \frac{R}{r_0^2} = 0, \quad R = e \pm i\tau/r_0$$

It is not possible to make a choice between these two solutions, as neither becomes infinite for $r = \infty$ (indeed, they both vanish, as a more accurate estimate will show). This already shows the greater manifold of solutions compared with the case (a) and expresses itself in the form of a continuous solution. Analogously to (a) we now set

$$\left. \begin{aligned} \rho &= 2r/r_0 = 2\sqrt{A}r, \quad R = e^{\pm \frac{i\rho}{2}}v, \\ R' &= e^{\pm \frac{i\rho}{2}}\left(v' \pm \frac{i}{2}v\right), \quad R'' = e^{\pm \frac{i\rho}{2}}\left(v'' \pm iv' - \frac{1}{4}v\right) \end{aligned} \right\} \quad (20)$$

and find for v the differential equation

$$v'' + \left(\frac{2}{\rho} \pm i\right)v' + \left[\left(\frac{B}{\sqrt{A}} \pm i\right)\frac{1}{\rho} - \frac{l(l+1)}{\rho^2}\right]v = 0 \quad (21)$$

which is analogous to (7). The assumption (7a) gives, as before, $\gamma = l$ as the only possible characteristic exponent at the zero point. But the "polynomial method" now fails on account of the imaginary coefficients in the differential equation. Nor would it lead to the desired end, since a polynomial, when multiplied by $e^{\pm i\rho/2}$, would become infinite for $\rho = \infty$. If we actually insert the assumption (7c) in (21) and equate the coefficients of ρ^{r+l-1} to zero, we get the recurrence formula

$$\left\{ \begin{aligned} &(\nu + l + 1)(\nu + l) + 2(\nu + l + 1) - l(l+1) \} a_{\nu+1} \\ &= \left\{ \mp i(\nu + l + 1) - \frac{B}{\sqrt{A}} \right\} a_{\nu} \end{aligned} \right\} \quad (22)$$

If we wished to force the series to end, we would obtain an *imaginary* value for $\frac{B}{\sqrt{A}}$, which would contradict our present assumption about A and E , respectively. *Consequently A and hence also E remains indeterminate.* We now have a *continuous spectrum* of proper values $E > 0$, which follows continuously on the limit of $E = 0$ of the discrete spectrum.

It is a particularly beautiful feature of Schrödinger's theory—and surprising from the mathematical point of view—that it connects the continuous spectrum by a uniform analytical process with the line-spectrum of hydrogen. The older theory (cf. *Atombau*, Chap. IX, § 7) required special though quite plausible physical assumptions to achieve this.

By (20) the radial component R of the proper function now becomes

$$R = e^{i\rho/2} \sum_0^{\infty} a_{\nu} \rho^{\nu+l} + e^{-i\rho/2} \sum_0^{\infty} a_{\nu}^* \rho^{\nu+l} \quad (23)$$

The a_{ν} 's are here to be calculated from (22); the fact that the coefficients in the second term of the summation are conjugate is directly clear by (22), if we choose a_0 to be real and equal in both terms. The appearance of the infinite series in (23) introduces no difficulty, and cannot affect the postulate that the proper function must remain finite for $\rho = \infty$. For we saw above that the two asymptotic particular solutions of R remain finite for $r = \infty$; but the particular solution implied in (23) which is finite at the zero point is linearly composed of these two. The present case is

different from that treated under (a), in which one of the two asymptotic particular solutions went to ∞ like $e + \rho^h$.

It is easy to estimate the asymptotic behaviour of R to a second approximation. For if we set

$$R = ae \pm i\rho^{1/2} \quad (23a)$$

and treat a , according to method given in eqns. (13c) of § 2, as a slowly variable quantity, we get from (3) the asymptotic differential equation

$$\frac{da}{d\rho} + \frac{1 \mp i\beta}{\rho} a = 0$$

with the abbreviation

$$\beta = Br_0 = \frac{B}{\sqrt{A}}$$

and obtain from it by integration

$$\log a = - (1 \mp \beta) \log \rho + \text{const.}$$

$$a = \frac{\text{const.}}{\rho} e^{\pm i\beta \log \rho} \quad (23b)$$

So we see: infinity is, indeed, an *essential singularity* for the proper function R , but if we move towards this point along the real axis, R certainly does not become infinite; rather it approaches the value *zero* with decreasing oscillations, as already remarked above.

Reverting to the original variable r we write the two asymptotic solutions contained in (23a and b) in the form

$$R = \frac{C}{r} e^{\pm i(kr + \alpha)} \quad (24)$$

which is usual for a spherical wave (moving outwards or inwards). In it we have set

$$k = \frac{1}{r_0}, \quad \alpha = \beta \log r \quad (24a)$$

C is a complex constant, α a slowly variable real quantity.

We shall require this method of representation later in a more rigorous form to explain the photo-electric effect.

We have here limited ourselves to prove the existence of the continuous spectrum and of its proper functions. This proof, in spite of its elementary character, seems to us convincing. We have not needed the more advanced analytical devices which Schrödinger calls into action for the continuous and the line spectrum, namely complex integrals and Laplace's differential equation. But we shall explain these methods in Note I. at the end of the book.

B. Two- and One- Dimensional Kepler Problem

If we treat eqn. (1) in plane co-ordinates r, ϕ , that is, set

$$\Delta\psi = \frac{\partial^2\psi}{\partial r^2} + \frac{1}{r} \frac{\partial\psi}{\partial r} + \frac{1}{r^2} \frac{\partial^2\psi}{\partial\phi^2},$$

we write the solution in the form

$$\psi = R e^{il\phi};$$

l is an integer, R a function of r alone. We get as the differential equation for R

$$\frac{d^2 R}{dr^2} + \frac{1}{r} \frac{dR}{dr} + \left(A + 2\frac{B}{r} + \frac{C}{r^2} \right) R = 0 \quad (25)$$

A and B have the same meaning as in (3a), but C is equal to $-l^2$.

In the case $E < 0$ we introduce the constant r_0 by means of (4) and the variable ρ by means of (4a). Since the asymptotic behaviour of (25) is the same as that of (3), v , too, is defined by the earlier eqn. (5). In place of (7) we then get as the differential equation for v

$$v'' + \left(\frac{1}{\rho} - 1 \right) v' + \left[\left(-\frac{B}{\sqrt{-A}} - \frac{1}{2} \right) \frac{1}{\rho} - \frac{l^2}{\rho^2} \right] v = 0.$$

With the assumption (7a) we find as the characteristic equation for the zero point: $\gamma(\gamma - 1) + \gamma - l^2 = 0$, and from this we get again $\gamma = +l$. The recurrence formula comes out as

$$\{ \quad \} a_{v+1} = \left\{ v + l + \frac{1}{2} - \frac{B}{\sqrt{-A}} \right\} a_v,$$

the factor of a_{v+1} having the same form as in the earlier recurrence formula (8). If the expansion is to cease at the term $v = n_r$, we must now set

$$\frac{B}{\sqrt{-A}} = n_r + l + \frac{1}{2},$$

from which we calculate

$$E = - \frac{2\pi^2 m e^4 Z^2}{h^2 (n_r + l + \frac{1}{2})^2} \quad (26)$$

The form of the Balmer term, including the Rydberg constant, thus comes out correctly also in the case of the plane Kepler problem; but the quantum number in the denominator is *half-integral* and not *integral*. Experiment, however, demands *integral* values. From this we conclude as has already been stated in § 3, C: whereas in the older theory (I, Chap. IV, §§ 1 and 7) two, one, or three dimensions could be used at will in the Kepler problem, we must use the full number of dimensions, three, in treating the problem by wave-mechanics. We must not avoid degenerate problems, as was formerly advised by Bohr, but must seek the highest degree of degeneracy, that is, here, the three-dimensional problem, in applying wave-mechanics and also quantum-mechanics.

It is interesting also to consider the one-dimensional Kepler problem. By this we mean the assumption of *one* co-ordinate in the wave-equation,

for example, the angular co-ordinate ϕ (corresponding to the circular orbit of the older theory.* Eqn. (1) then assumes the form

$$\frac{1}{r^2} \frac{\partial^2 \psi}{\partial \phi^2} + \frac{8\pi^2 m}{h^2} \left(E + \frac{e^2 Z}{r} \right) \psi = 0.$$

Its solution, periodic in ϕ , is

$$\psi = e^{in\phi}$$

if we take n as *integral* and make it subject to the condition

$$-\frac{n^2}{r^2} + \frac{8\pi^2 m}{h^2} \left(E + \frac{e^2 Z}{r} \right) = 0 \quad . \quad . \quad . \quad (27)$$

We are troubled with the indefinite denominator r . We can eliminate it by making use of the rule of the older quantum theory that the mean value of the potential energy is equal to twice the total energy. We shall elucidate in Chap. II, § 9, what this rule signifies in wave-mechanics and on what it is based. In our case, for which r is constant, it states simply that

$$-\frac{e^2 Z}{r} = 2E, \quad r = -\frac{e^2 Z}{2E}.$$

If we substitute this in (27), we get

$$\frac{4E^2 n^2}{e^4 Z^2} + \frac{8\pi^2 m E}{h^2} = 0,$$

$$\text{that is,} \quad E = -\frac{2\pi^2 m e^4 Z^2}{n^2 h^2} \quad . \quad . \quad . \quad (28)$$

This is again the Balmer term, and in this case n is *integral*. We thus find the same remarkable change from integral to half-integral and back again to integral quantum numbers in passing from one dimension to two and three dimensions, as we found earlier in § 3, C, for the oscillator, and likewise for the rotator.

C. *Numerical and Graphical Representations of Proper Functions. Meaning of the Quantum Numbers in Wave-Mechanics. Comparison with the Earlier Ideas involving Orbits. Spherical and Axial Symmetry of the Charge Cloud.*

We next give a tabular summary of the proper functions of the Kepler problem in the case of discrete states. As a preliminary step we represent in Fig. 5 the series of Laguerre polynomials; the expressions given in the figure follow directly from (14). The general formula is

$$L_n(x) = (-1)^n \left(x^n - \frac{n^2}{1!} x^{n-1} + \frac{n^2(n-1)^2}{2!} x^{n-2} + \dots + (-1)^n n! \right) \quad (29)$$

* We might also introduce elliptic co-ordinates and, as an analogy to the periodic elliptic orbit, work out a solution of the wave-equation which depends periodically on one of the two elliptic co-ordinates.

To explain Table 1 on p. 71, we further remark that in the

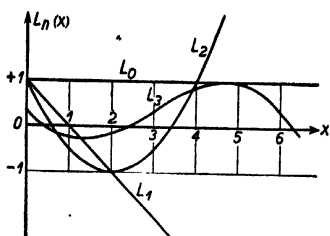


FIG. 5.—The Laguerre polynomials:

$$L_0 = 1.$$

$$L_1 = -x + 1.$$

$$L_2 = x^2 - 4x + 2.$$

$$L_3 = -x^3 + 9x^2 - 18x + 6.$$

In the figure L_2 and L_3 have been multiplied by the factors $\frac{1}{2}$ and $\frac{1}{6(1+\sqrt{3})}$, which have been chosen thus in order that the principal maxima (or minima) of L_2 and L_3 become equal to ± 1 .

expression P_l^m of the spherical harmonic $l = 0$ makes $m = 0$, since we always have

$$|m| \leq l,$$

and that it follows from $n = 1$ that $l = 0$ and $n_r = 0$, since, by (9a), the relation

$$l + n_r = n - 1$$

holds.

We then calculate the length r_0 defined by (4). By taking (3a) and (10) into consideration we get

$$r_0^2 = \frac{h^2}{8\pi^2 m} \cdot \frac{h^2 n^2}{2\pi^2 m e^4 Z^2},$$

$$r_0 = \frac{h^2}{4\pi^2 m e^2} \cdot \frac{n}{Z} = \frac{a n}{Z} \quad (30)$$

Here a is the radius of the first Bohr circle in the older theory of the hydrogen atom (cf. I, Chap. IV, § 3, eqn. 7), and so a/Z is the corresponding radius for a nucleus with charge Ze . So our reflections based on the theory of functions, in particular our method of asymptotic approximation, which led to the definition of the quantity r_0 , point directly to this orbital radius a .

If we measure r in units of the radius a/Z , that is, set

$$s = Z \frac{r}{a} \quad (30a)$$

our definition (4a) for the variable ρ becomes

$$\rho = \frac{2Z}{n} \frac{r}{a} = \frac{2}{n} s \quad (30b)$$

Concerning the values of P_l^m we refer to Fig. 2, § 2, and concerning the expressions for R to eqns. (13) and (18a). For example, for $n = 1$ and $l = 0$ we have

$$R N_r = e^{-\rho/2} \frac{d}{d\rho} L_1 = -e^{-\rho/2}, \quad \frac{1}{N_r} = \pm 2 \left(\frac{Z}{a} \right)^{3/2}.$$

On account of the choice of sign in the case of N_r we can always arrange that ψ , for example, becomes positive for $r = 0$. The value of ψ in the last column but one of the table then follows from eqn. (2). The last column contains the usual spectroscopic notation of terms and the

TABLE 1

n	L	m	n_r	P_l^m	$R(\rho)$	$-\frac{1}{2}\left(\frac{a}{2}\right)^{3/2} \cdot \frac{R(\rho)}{N_r}$	ψ normalised $\sqrt{\pi\left(\frac{a}{2}\right)^{3/2}}$	Terms and Shells
1	0	0	0	1	$-e^{-\rho/2}$	e^{-s}	e^{-s}	$1s, K$
2	0	0	1	1	$2(\rho-2)e^{-\rho/2}$	$\frac{1}{4\sqrt{2}}(2-s)e^{-s/2}$	$\frac{1}{4\sqrt{2}}(2-s)e^{-s/2}$	$2s, L_I$
2	1	0	0	$\cos \theta$	$-6\rho e^{-\rho/2}$	$\frac{1}{4\sqrt{6}}e^{-s}$	$\frac{1}{4\sqrt{2}}e^{-s/2} \cos \theta$	$\left. \begin{array}{l} 3p, L_{II} + L_{III} \end{array} \right\}$
2	1	1	0	$\sin \theta$	$-6\rho e^{-\rho/2}$	$\frac{1}{4\sqrt{6}}e^{-s}$	$\frac{1}{8}e^{-s/2} \sin \theta e^{i\phi}$	
3	0	0	2	1	$-3(\rho^2-6\rho+6)e^{-\rho/2}$	$\frac{1}{81\sqrt{3}}(27-18s+2s^2)e^{-s/2}$	$\frac{1}{81\sqrt{3}}(27-18s+2s^2)e^{-s/2}$	$3s, M_I$
3	1	0	1	$\cos \theta$	$24\rho(\rho-4)e^{-\rho/2}$	$\frac{2}{81\sqrt{6}}s(6-s)e^{-s/2}$	$\frac{2s}{81\sqrt{2}}(6-s)e^{-s/2} \cos \theta$	$\left. \begin{array}{l} 3p, M_{II} + M_{III} \end{array} \right\}$
3	1	1	1	$\sin \theta$	$24\rho(\rho-4)e^{-\rho/2}$	$\frac{2}{81\sqrt{6}}s(6-s)e^{-s/2}$	$\frac{s}{81}(6-s)e^{-s/2} \sin \theta e^{i\phi}$	
3	2	0	0	$\frac{3}{2} \cos^2 \theta - \frac{1}{2}$	$-120\rho^2 e^{-\rho/2}$	$\frac{2}{81\sqrt{30}}e^{-s}$	$\frac{2s^2}{81\sqrt{6}}e^{-s/2} \left(\frac{3}{2} \cos^2 \theta - \frac{1}{2} \right)$	$3d, M_{IV} + M_V$

allocation to the shells of the fully occupied atom (cf. *Atombau*, Chap. IV, p. 280, Table XVIII), which we express in Bohr's notation. The fact that according to this last column in many instances *two* different shells correspond to *one* hydrogen state is due to our not having hitherto taken into account the differentiation of state due to the "spinning electron."

In Fig. 6 we represent graphically the radial part of the proper function ψ . The following relation of great generality is shown here: the number of zero points between the limiting points $s = 0$ and $s = \infty$ is exactly equal to the radial quantum number n_r . The curves K, $L_{II} + L_{III}$ and $M_{IV} + M_V$ in Fig. 6 have no zero point; by our Table they correspond to the case $n_r = 0$. In the case of L_I and also in that of $M_{II} + M_{III}$ there is one zero point; for then, by our Table, $n_r = 1$, and the zero points in question lie at $s = 2$ and $s = 6$ respectively. At M_I , two zero points occur, corresponding to $n_r = 2$; according to the Table they are given by $2s^2 - 18s + 27 = 0$ and so lie at $s = \frac{3}{2}(3 \pm \sqrt{3})$.

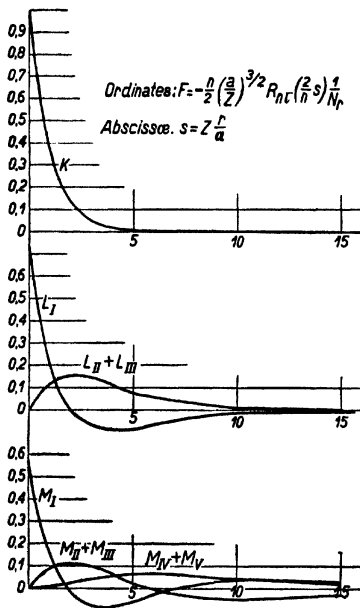


FIG. 6.

The curves represent the normalised radial part of ψ : $R_{nc}(\frac{2}{n}) \cdot \frac{1}{N_r}$, multiplied by the factor $-\frac{n}{2}(\frac{a}{Z})^{3/2}$. The factor n makes the ordinates of the curves L have twice the value that they are entitled to have according to normalisation, and makes the curves M correspondingly appear three times higher.

series, that is, by the degree of the resulting polynomial: *quantum numbers denote the numbers of nodes in the proper function that lie between the limiting points for the co-ordinate in question.* This brings to mind the analogy of a vibrating string in which the ordinal number of an overtone is likewise measured by the number of nodes that lie between the fixed ends of the string.

This leads to a simple wave-mechanical interpretation of the quantum numbers; indeed, not only in the case of the radial co-ordinate and of the Kepler problem but in all cases where, using the polynomial method, we can apply the following definition by forcibly breaking off an expansion in

According to O. Perron the proof may be carried out as follows. Let P_0, P_1, \dots, P_n be a system of orthogonal polynomials belonging to the interval (a, b) and of degree $0, 1, \dots, n$, such that for every $m < n$ we have

$$\int_a^b P_m(x) P_n(x) p(x) dx = 0 \quad (31)$$

Here p denotes an arbitrary "weight-function" which does not vanish for $a < x < b$, and may be taken e.g. as positive; the P 's can, indeed, be regarded as defined by (31), in which case the numerical factor would have to be fixed by a corresponding normalising condition. Let us now assume that P_n has only $\nu < n$ roots x_1, x_2, \dots, x_ν between a and b , and let us form

$$g_\nu(x) = (x - x_1)(x - x_2) \dots (x - x_\nu)$$

Then $g_\nu(x) P_n(x)$ is a function of uniform sign in the interval a to b , and the inequality

$$\int_a^b g_\nu(x) P_n(x) p(x) dx \neq 0 \quad (31a)$$

certainly holds. On the other hand g_ν may be compounded linearly from the polynomials P_0, P_1, \dots, P_ν :

$$g_\nu(x) = \sum_0^\nu c_\mu P_\mu.$$

Hence by (31) the integral in (31a) would have to become *equal* to zero. Consequently $\nu < n$ is impossible.

$\nu = n$ leads to no contradiction, as then one member of the sum contains the normalising condition in place of the condition for orthogonality. Since $\nu > n$ is excluded by the fundamental theorem of algebra, it necessarily follows that *every polynomial P_n of an orthogonal system has exactly n zero points between a and b .* If in accordance with our above remarks we identify the degree n of the polynomial with the quantum number, this number becomes equal to the number of zero points of the polynomial or, as we said before, equal to the number of "nodes."

Reverting to the Kepler problem we add that the weight-function p just used has the significance

$$p(\rho) = \rho^{2(l+1)} e^{-\rho}$$

in the case of our derived Laguerre polynomial, eqn. (13), as can be seen from the differential eqn. (11). Our theorem also holds, of course, for the azimuthal quantum number l , which in the case of $m = 0$ specifies the degree of the Legendre polynomial P . The fact that this polynomial has exactly l zero points between $\cos \theta = -1$ and $+1$ has long been known and manifests itself in Fig. 1 of p. 12: $P_1 = \cos \theta$ vanishes *once*, $P_2 = \frac{3}{2} \cos^2 \theta - \frac{1}{2}$ vanishes *twice*. In the case $m \neq 0$ the degree of the polynomial belonging to P_l^m (we called it P_{l-m}^m in § 6) becomes equal to $l - m$. Hence the function $P_1^1 = \sin \theta$ that occurs in our table has *no* zero point within the interval under consideration.

From the proper functions ψ we pass on to the associated charge densities $e\psi\psi^*$,* which we originally defined in eqn. (16) of § 5. We first deal with the general symmetry of this distribution of charge. We then see that in the case of the S-terms, that is, for $l = 0$, the distribution is *spherically symmetrical*, that is, independent of θ and ϕ ; in the case of the P, D . . . terms, that is, for $l > 0$, no matter whether $m = 0$ or > 0 , the distribution becomes axially symmetrical about the polar axis $\theta = 0$, which however is only mathematically distinguished. Concerning the last remark we must add that it rests on the particular form $e \pm im\phi$ which expresses the dependence on ϕ , and in which $\psi\psi^*$ becomes independent of ϕ . In the case of the more general assumption *

$$ae + im\phi + be - im\phi,$$

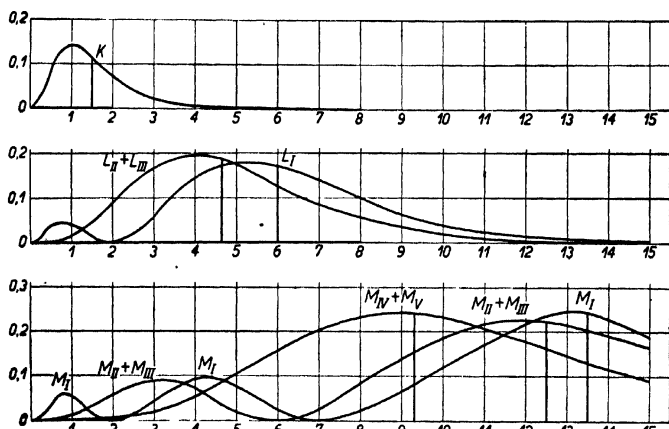


FIG. 7.—Graphical representation of the distribution of charge in the various proper states. The function s^2F^2 is plotted as the ordinate, F having the same value as in Fig. 6. The L-curves accordingly have their ordinates multiplied four times, the M-curves their nine times (cf. the remarks attached to Fig. 6).

the axial symmetry is obviously replaced by an axial periodicity. Further we must note that the position of the polar axis (when no magnetic field is acting) is physically indefinite, so that the symmetry or periodicity in question can occur about any axis in space. This clearly corresponds to the spatially indefinite orientation of the orbital planes in the earlier theory.

In Fig. 7,† however, we shall not plot the charge-density itself, but in

* On account of the normalisation the constants a and b here introduced are bound by the condition

$$|a|^2 + |b|^2 = 1.$$

† These curves, like those in Fig. 6, are in the main taken from a paper by L. Pauling, Proc. Roy. Soc., 114, 181 (1927).

each case the charge integrated over the spherical surface of radius s [eqn. (30a)], namely

$$\bar{\rho} = s^2 \frac{R^2}{N_r^2} \quad (32)$$

The resulting curves are instructive in several respects. Firstly, they illustrate how far the "charge-cloud" (*Ladungswolke*) into which the electronic charge e becomes merged according to Schrödinger (cf. also § 8), extends itself. The fact that $\bar{\rho}$ vanishes for $s = 0$ whereas ψ itself had a maximum there in several of the curves of Fig. 6, is, of course, due to the factor s^2 in (32). Starting from $s = 0$, $\bar{\rho}$ either increases directly to a maximum or first passes through some smaller maxima to a *principal maximum*, after which it decreases exponentially as far as infinite values of the abscissa.

We compare this distribution of charge with the shells of the corresponding n_k -orbits of the older theory. According to the new view these shells are by no means thin, sharply defined shells, but very much widened regions of continuous space charge, which may even be subdivided regions. But for each of these distributions of charge we may define a sphere of *mean radius* s_m such that it divides the total charge into two equal parts. The corresponding ordinate of the charge-distribution is exhibited in the figure and in each case lies near the principal maximum. If we call this s_m for brevity the radius of the shell, we see that the radius of the K-shell is the smallest, the two radii of the L-shell are greater, and those of the M-shell are greatest, quite in accord with the earlier ideas.

There is a still further correspondence between the earlier ideas and our present view. To illustrate this the dimensions of the hydrogen orbits according to the older theory are shown in the figure, namely for the K and L_I shells (circular orbits): they are marked off along the x -axis as dark lines and extend from $s = 0$ to the value of s that corresponds to the circular radius in question. *We see that the dimensions of the paths of the earlier theory coincide in the main with those of the present charge-distribution.*

D. Motion of the Nucleus

Hitherto we have treated the Kepler problem as a one-body problem. We now pass on to the two-body problem. We may deal with it very quickly by referring to § 3, E. We start out, as in § 3, E, from the partial differential equation (23) in the phase space of six dimensions, namely in the space of the co-ordinates $x_1 y_1 z_1$ of the electron and $x_2 y_2 z_2$ of the nucleus, and build up from them the co-ordinates $\xi \eta \zeta$ of the centre of inertia and the relative co-ordinates xyz . The wave-function Ψ of the

whole problem splits up into the wave-function ψ of the "relative motion" and χ of the "motion of the centre of inertia":

$$\left. \begin{aligned} \Psi(x_1 y_1 z_1 x_2 y_2 z_2) &= \psi(x y z) \chi(\xi \eta \zeta), \\ x &= x_1 - x_2, \dots (m + M)\xi = m x_1 + M x_2, \dots \end{aligned} \right\} \quad (33)$$

in which we have called the masses of the electron and the nucleus, as usual, m and M . The differential equations (27) and (26) of § 3 hold for ψ and χ , the significance of the potential energy V being correspondingly altered. We write the equation (24) defining the "resultant mass" μ in our present nomenclature thus

$$\frac{1}{\mu} = \frac{1}{m} + \frac{1}{M}, \quad \mu = \frac{mM}{M + m} = \frac{m}{1 + m/M}.$$

Eqn. (26) of § 3 now, too, possesses the continuous spectrum there described (cf. also the end of § 1), corresponding to the indefiniteness of the motion of the centre of inertia, or the arbitrariness in the choice of the system of reference. On the other hand, eqn. (27) of § 3 is identical with our eqn. (1) of the present section, except that m is replaced by μ and E by $E - E_t$ (E_t = energy of the motion of translation). Consequently, everything that we have derived about proper values and proper functions can be taken over from the one-body problem and applied to the two-body problem. In particular, the spectral equation (10) remains in force, *except that m is replaced by μ* . The *Rydberg constant* R accordingly now has the value

$$R = \frac{2\pi^2 \mu e^4}{h^3} = \frac{2\pi^2 m e^4}{h^3 (1 + m/M)} \quad (34)$$

It is unnecessary here to do more than point out the beautiful way in which this formula has been confirmed spectroscopically (by the small distances separating the Balmer and the Pickering lines, I, Chap., IV, p. 222). The new theory accounts just as well for this and by means of the same formulæ as the old theory. But what are we to say about the assertion on p. 223 (of the same chapter) that "the differences between the lines (namely of the Balmer and Pickering series) give us definite information to the effect that in our intra-atomic planetary system the law concerning the persistence of the common centre of gravity remains in force"? Can we still maintain this assertion? For we have now no planetary system, no concentrated masses or charges, still less in the case of the nucleus than in that of the electron (cf. below). Yet there can be no doubt that our result rests on the law of the persistence of the centre of inertia of mechanics. It retains its place also in micro-mechanics, even if it does not here refer to point-masses but to *mean values* (cf. Chap. II, § 9). It is not the certainty of mathematical statements but only the vividness of the mental pictures (*Anschaulichkeit der Vorstellung*)

that becomes weakened in micro-mechanics as compared with macro-mechanics.

We shall now answer the question as to what "charge-cloud" is to be ascribed to the nucleus in the two-bodies problem; the charge-cloud of the electron relative to the nucleus will remain essentially the same as in the one-body problem. We shall seize this opportunity to explain the rule* given by Schrödinger for defining in the case of several point charges the charge distribution to be ascribed to each individually. This rule, specialised for our two-bodies problem, runs: keep the co-ordinates x_2, y_2, z_2 of the nucleus fixed and integrate the absolute value of the proper function Ψ over all possible co-ordinates x_1, y_1, z_1 of the electron. Multiplying the result by e , we get the density ρ_2 of the charge-cloud of the nucleus at the point x_2, y_2, z_2 . By eqn. (33) this gives in our case

$$\begin{aligned}\rho_2 &= e \iiint dx_1 dy_1 dz_1 |\Psi|^2 \\ &= e \iiint dx_1 dy_1 dz_1 |\psi(xyz)|^2 |\chi(\xi\eta\zeta)|^2 \quad . \quad . \quad . \quad (35)\end{aligned}$$

Instead of $x_1 \dots$ we may now introduce $x = x_1 - x_2$ as the variable of integration, where, by (33),

$$\xi = x_2 + \frac{m}{m+M}x, \dots$$

So if we neglect the ratio $\frac{m}{M}$, $\xi \dots$ and hence also χ becomes independent of the variables of integration x, y, z . In place of (35) we may therefore write

$$\rho_2 = e |\chi(\xi, \eta, \zeta)|^2 \iiint dx dy dz |\psi|^2.$$

But the integral that remains in this is equal to unity by the normalising condition, and we therefore get

$$\rho_2 = e |\chi(\xi, \eta, \zeta)|^2.$$

Now, we had $\chi = e^{ikx}$ except for a constant (cf. E, § 3), thus $|\chi| = \text{const.}$ Accordingly also $\rho_2 = \text{const.}$

The charge-cloud of the nucleus is uniformly distributed over the whole of space to an infinite distance, and of course with the density zero, since we must postulate

$$\int \rho_2 d\tau_2 = e.$$

In essence this asserts nothing more than that in the absence of external fields every point of space is equally probable as the centre of inertia of an atom. If the atom is enclosed in a cavity (*Hohlraum*), the remarks at the conclusion of E, § 3, must be observed: the density of the charge-cloud of the nucleus becomes finite but is nevertheless still appreciably constant in the whole of the cavity.

E. The Selection Rules for the Kepler Problem

The quantities $q_{n,m}$ ["matrix elements," eqn. (21), § 5], which lead to the selection rules, must now be written with six indices. Combining

* Schrödinger, *Collected Papers on Wave Mechanics, Quantisation as a Problem of Proper Values*, Part IV, §§ 2 and 7 (Blackie & Son, Ltd.).

the co-ordinates into complex quantities as in § 6 and calling them $\xi\eta\zeta$, we obtain

$$\left. \begin{aligned} (\xi + j\eta)_{nlm, n'l'm'} &= \int r \sin \theta e^{j\phi} \psi_{nlm} \psi_{n'l'm'}^* d\tau = [r][\theta\phi]_1, \\ \zeta_{nlm, n'l'm'} &= \int r \cos \theta \psi_{nlm} \psi_{n'l'm'}^* d\tau = [r][\theta\phi]_2. \end{aligned} \right\} \quad (36)$$

The (θ, ϕ) -components of these quantities agree exactly with the corresponding quantities in the case of the rotator in B, § 6, except for the symbols of the quantum numbers (we now feel the compulsion to write l, m instead of the earlier m, μ as we must adapt our present notation to the requirements of series spectra, whereas the earlier symbols were chosen to meet the needs of band spectra). Consequently we can take over the selection rules for our present quantum numbers l, m directly from the previous section, eqns. (33) and (32). *Only the following transitions are allowed:*

$$l \rightarrow l \pm 1 \quad . \quad . \quad (37) \qquad m \begin{matrix} \nearrow m \\ \searrow m \pm 1 \end{matrix} \quad . \quad . \quad (37a)$$

(37) is the selection rule for the azimuthal quantum number, which governs the whole theory of series spectra; (37a) is the selection rule for the equatorial or magnetic quantum number, which becomes effective when the lines are resolved magnetically.

We then come to the radial component $[r]$; this is the same in both eqns. (36), namely

$$[r] = \frac{1}{N_{nl}} \frac{1}{N_{n'l'}} \int_0^\infty r^3 R_{nl} R_{n'l'} dr \quad . \quad . \quad (38)$$

in which, by (37), l' can only have the two values $l' = l \pm 1$ and in which we now have to denote the normalising factor N_r , used earlier, by N_{nl} or $N_{n'l'}$. The question is whether there is also a selection rule for this radial component, which limits the transitions of the radial quantum number n_r or, what is the same thing, limits the "principal quantum number" n . We know that this is not the case, as otherwise there would be no series spectra. But we must endeavour to understand this from the form of the expression (38).

By (30b) the arguments of R_{nl} and $R_{n'l'}$ themselves depend on n and n' , respectively; for they are

$$\frac{2}{n} s \quad \text{and} \quad \frac{2}{n'} s \quad \text{respectively, where } s = Zr/a.$$

Written more fully, and with due regard to the notation of (13), (38) becomes:

$$\left. \begin{aligned} [r] &= C \cdot S, \quad C = \frac{1}{N_{nl}} \cdot \frac{1}{N_{n'l'}} \left(\frac{a}{Z}\right)^4 \left(\frac{2}{n}\right)^l \left(\frac{2}{n'}\right)^{l'}, \\ S &= \int_0^\infty s^{l+l'+1} L_{n+l}^{2l+1} \left(\frac{2}{n}s\right) L_{n'+l'}^{2l'+1} \left(\frac{2}{n'}s\right) e^{-\left(\frac{1}{n}+\frac{1}{n'}\right)s} ds \end{aligned} \right\} \quad (38a)$$

We shall now endeavour to apply the simple argument of § 6, A, by comparing the integral S with the condition of orthogonality of the radial proper functions. By (16), the same variable of integration s being used, this condition is

$$\int_0^\infty L_{n'+l}^{2l+1}\left(\frac{2}{n}s\right)L_{n+l}^{2l+1}\left(\frac{2}{n}s\right)p(s)ds = 0 \quad (38b)$$

in which the "weight-function" $p(s)$ has the significance

$$p(s) = s^{2l+2}e^{-\left(n+\frac{1}{n}\right)s} \quad (38c)$$

For this purpose we combine the first two factors under the integral sign in S, after splitting off the factor s^{2l+2} , to a polynomial G (of degree $n'-l$) and write S as follows:

$$S = \int_0^\infty G(s)L_{n'+l}^{2l+1}\left(\frac{2}{n}s\right)p(s)ds \quad (38d)$$

Of course we can now also develop G in terms of the polynomials L:

$$G = \sum_{\nu=0}^{n'-l} c_\nu L_{\nu+2l+1}^{2l+1} \quad (39)$$

But the condition of orthogonality fails us, as it no longer enables us to carry out the integration (38d).

For if, as a result of the comparison with (38b), we make $\nu + 2l + 1 = n' + l$, that is, $n' = \nu + l + 1$, p becomes dependent on ν , whereas in (38d) p is to have the same value (38c) for all members of the sum (39). Consequently the condition of orthogonality does not here lead to the calculation of $[r]$ and, in particular, gives rise to *no selection rule for the radial quantum number*.

F. Questions of Intensity. Lyman and Balmer Series

The circumstances just described also increase the difficulty of calculating intensities generally. We therefore restrict ourselves to the simplest cases.

(a) In the Lyman series the final state is $n = 1$, and hence certainly $l = 0$. In the initial state n' is arbitrary, but by (37) l' certainly = 1. So we have

$$\left. \begin{aligned} L_{n'+1}^{2l+1}\left(\frac{2}{n}s\right) &= L_1^1(2s) = -1 \text{ [eqn. (29)],} \\ L_{n'+1}^{2l'+1}\left(\frac{2}{n}s\right) &= L_{n'+1}^3\left(\frac{2}{n}s\right) \end{aligned} \right\} \quad (40)$$

If we choose $2s/n' = x$ as a new variable of integration and, from now on, omit the dash from n' then, by (38a), we have

$$S = -\left(\frac{n}{2}\right)^3 \int_0^\infty x^4 L_{n+1}^3(x) e^{-\frac{n+1}{2}x} dx \quad (40a)$$

To perform this integration we use a special device. In the exponential function we temporarily replace $\frac{1}{2}(n+1)$ by an arbitrary number α and write

$$S = - \left(\frac{n}{2}\right)^5 \frac{d^4}{d\alpha^4} S_1, \quad S_1 = \int_0^\infty e^{-\alpha x} \frac{d^3}{dx^3} L_{n+1}(x) dx. \quad (40b)$$

Integrating by parts three times, in which process we may omit the terms independent of the integral sign, as they will drop out of themselves in the following differentiations with respect to α , we get, taking (14) into account:

$$S_1 = \alpha^3 \int_0^\infty e^{-\alpha x} L_{n+1}(x) dx = \alpha^3 \int_0^\infty e^{-(\alpha-1)x} \frac{d^{n+1}}{dx^{n+1}} (x^{n+1} e^{-x})$$

Further integrations by parts lead to

$$S_1 = \alpha^3 (\alpha-1)^{n+1} \int_0^\infty e^{-\alpha x} x^{n+1} dx.$$

Introducing $y = \alpha x$ as a new variable of integration,

$$S_1 = \frac{(\alpha-1)^{n+1}}{\alpha^{n+1}} \int_0^\infty e^{-y} y^{n+1} dy = \frac{(\alpha-1)^{n+1}}{\alpha^{n+1}} (n+1)!,$$

and so, by (40b)

$$S = - \left(\frac{n}{2}\right)^5 (n+1)! \frac{d^4 A}{d\alpha^4}, \quad A = \frac{(\alpha-1)^{n+1}}{\alpha^{n+1}}. \quad (40c)$$

We next differentiate A successively with respect to α . After three differentiations the factor

$$(n+1)n(n-1) = \frac{(n+1)!}{(n-2)!}$$

may be extracted. Four differentiations give

$$A^{IV} = \frac{(n+1)!}{(n-2)!} \frac{\left(1 - \frac{1}{\alpha}\right)^{n-3}}{\alpha^6} (n+2-4\alpha);$$

if we now write for α its value $\frac{1}{2}(n+1)$,

$$A^{IV} = -2^6 n \frac{(n+1)! (n-1)^{n-3}}{(n-2)! (n+1)^{n+3}}$$

and so, by (40c),

$$S = +2n^6 \frac{[(n+1)!]^2 (n-1)^{n-3}}{(n-2)! (n+1)^{n+3}}.$$

In our case (for which we set $n=1$, $l=0$ and $n'=n$, $l'=1$) the quantity C in (38a) becomes

$$C = \frac{1}{N_{1,0}} \frac{1}{N_{n,1}} \left(\frac{a}{Z}\right)^4 \cdot \frac{2}{n}.$$

Hence by (38a)

$$[r] = 4 \frac{1}{N_{1,0}} \frac{1}{N_{n,1}} n^6 \left(\frac{a}{Z}\right)^4 \frac{[(n+1)!]^2 (n-1)^{n-3}}{(n-2)! (n+1)^{n+3}}. \quad (41)$$

To calculate the N 's we must use eqn. (18a). For $n = 1$, $l = 0$ we also have $m = 0$ and $r_0 = a/Z$ by (30). So that for the final term we get

$$\frac{1}{N_{1,0}^2} = 4 \cdot \left(\frac{Z}{a}\right)^3.$$

If we restrict ourselves to the ζ -component of the emission (the ξ - and η -components must yield the same final result), then, on account of the selection rule (37a), we must also set $m = 0$ for the initial term. Eqn. (18a) then gives

$$\frac{1}{N_{n,1}^2} = 4 \left(\frac{Z}{an}\right)^3 \frac{(n-2)!}{n[(n+1)!]^3}.$$

Consequently

$$\frac{1}{N_{1,0}} \frac{1}{N_{n,1}} = 4 \left(\frac{Z}{a}\right)^3 \sqrt{\frac{(n-2)!}{(n+1)!}} \frac{1}{n^2(n+1)!}$$

Substituting in (41),

$$[r] = 16 \left(\frac{a}{Z}\right) n^3 \sqrt{\frac{(n+1)!}{(n-2)!}} \frac{(n-1)^{n-3}}{(n+1)^{n+3}} \quad (41a)$$

By the second eqn. (36) this expression is yet to be multiplied by the factor $[\theta\phi]_2$, which, by § 6, comes out as $\frac{1}{\sqrt{3}}$.

To obtain a measure of the intensity we must remember that our matrix elements ζ , ξ , η signify electric moments M_z , M_x , M_y with the time factor $e^{i\omega t}$, where ω is the frequency of the corresponding spectral line. The intensity observed in any particular direction, for example the positive direction of z , is calculated from the total moment M by differentiating twice with respect to t (cf., for example, I, Chap. I, p. 25, eqn. (2)) in accordance with the formula

$$J = \frac{\dot{M}^2 \sin^2 \theta}{4\pi c^3} = \frac{\dot{M}_x^2 + \dot{M}_y^2}{4\pi c^3} = \frac{2\dot{M}_z^2}{4\pi c^3} = \frac{\omega^4 M_z^2}{2\pi c^3}$$

θ denotes the angle between the vector \dot{M} and the direction of observation z .

Now, for the n^{th} line of the Lyman series

$$\frac{\omega}{2\pi c} = R \left(1 - \frac{1}{n^2}\right), \text{ that is, } \frac{\omega^4}{16\pi^4 c^4 R^4} = \frac{(n-1)^4(n+1)^4}{n^8}.$$

We therefore get for the intensity *

$$J = \frac{2^7 (n+1)! (n-1)^{2n-2}}{n^2 (n-2)! (n+1)^{2n+2}}$$

or, written more simply,

$$J = \frac{2^7 (n-1)^{2n-1}}{n(n+1)^{2n+1}} \quad (42)$$

* We have cancelled the universal factor $\frac{2^4 \pi^7 m^2 e^{12}}{3 c^5 h^3 Z^3}$, in order to get into agreement with Pauli's result (cf. next footnote and Schrödinger's *Collected Papers*, p. 101).

Asymptotically for $n \rightarrow \infty$ this gives

$$J \sim \frac{1}{n^3}$$

as can also be deduced from the correspondence principle for any arbitrary series.

(b) In the Balmer series the final state is $n = 2$ and $l = 0$ or 1. In the first case the selection rules (37) give $l' = 1$ for the initial state, in the second case l' can be 2 or 0. So we have three possibilities which have already been kept distinct in I, p. 336, and have been characterised by the series notation: principal series, diffuse series (first subsidiary series), sharp series (second subsidiary series):

$n = 2, l = 0$ $n' > 2, l' = 1$ $n'p \rightarrow 2s$	$n = 2, l = 1$ $n' > 2, l' = 2$ $n'd \rightarrow 2p$	$n = 2, l = 1$ $n' > 2, l' = 0$ $n's \rightarrow 2p$
Principal series.	Diffuse series.	Sharp series.

Here, too, the intensities may be calculated by the method used for the Lyman series, but the working is a little more cumbersome.

We shall only quote the results.

1. Principal series	$J = \frac{2^9(n^2 - 1)(n - 2)^{2n-2}}{n(n + 2)^{2n+2}}$
2. Diffuse series	$J = \frac{2^{12}n(n^2 - 1)(n - 2)^{2n-3}}{3(n + 2)^{2n+3}}$
3. Sharp series	$J = \frac{2^7n(n - 2)^{2n-2}}{3(n + 2)^{2n+2}}$

As the sum of all three series, which alone is observed in the case of hydrogen, one gets

Total intensity

$$1. + 2. + 3. \quad J = \frac{2^7(n - 2)^{2n-3}}{n(n + 2)^{2n+2}}(15n^4 - 32n^2 + 16).$$

The formulæ for the intensity of the lines in the Lyman and the summed Balmer series were first calculated by W. Pauli and communicated by Schrödinger * together with a general method of representing the corresponding matrix elements as series. Sugiura † has extended the calculation to the Paschen series $R\left(\frac{1}{3^2} - \frac{1}{n^2}\right)$. The component series of the Balmer lines and a great number of higher series have been calculated by A. Kupper.‡ A knowledge of the component series is necessary for making a comparison with the observed series of the alkalis. This comparison comes out very favourably in the case of the absorption measurements of the principal series carried out by Trumphy,§

* *Collected Papers, Quantisation and Proper Values*, III, p. 101.

† *Journ. de phys.*, série 6, 8, 113 (1927); *ZS. f. Phys.*, 44, 190 (1927).

‡ *Diss. München, Ann. d. Phys.*, 86, 511 (1928).

§ *ZS. f. Phys.*, 42, 327 (1927); 44, 575 (1927).

if the various alkalis are rationally allocated to the different hydrogen series (Li to the Balmer series, Na to the Paschen series, K to the Brackett series, and so forth). On the other hand, the emission flame spectra measured by Miss Bleeker * do not agree at all with the theory, nor do the hydrogen series themselves, which seem particularly sensitive to the conditions of excitation.†

§ 8. Statistical View of the Continuous Charge-cloud. Complete and Incomplete Shells. Magnetic Moments. Spatial Quantising. Physical Applications (Sphere of Action in Collisions, Supposed Double Refraction in a Magnetic Field, Lattice Forces)

A. Statistical View of the Continuous Charge-cloud

We refuse from the outset to take literally the charge-cloud to which Schrödinger's theory leads. Rather we shall retain the well-founded view that the electron is point-like in form or at any rate is a configuration of sub-atomic dimensions. We are led to this belief by all experiments with cathode rays, their passage through metallic films, and the indivisibility of the electronic charge. A theoretical factor also enters, which is furnished by the Schrödinger theory itself: in the assumption of the wave-equation every electron is treated as a point in the phase space and calculations are made with definite co-ordinates x, y, z ; that is, the electron is not treated as a continuous region. For example, in the Kepler problem we set the potential energy between the nucleus and the electron proportional to $(x^2 + y^2 + z^2)^{-\frac{1}{2}}$ in the wave-equation, and we thus give the nucleus the definite co-ordinates 0, 0, 0, and the electron the likewise discrete co-ordinates x, y, z . If deductions from the wave-equation lead to an apparently opposite result, no literal meaning can be attached to it.

Consequently the charge-cloud can have only a *statistical* meaning. Giving up the idea of individual orbits we regard the charge-cloud as the sum-total of possible paths of the electron and imagine the average time of stay of the electron in each individual position as determined by the charge-density at that point. This statistical view, first given a logical basis in the papers of Born,‡ was placed on broader foundations by Pauli,§ Jordan,|| Heisenberg,** Dirac,†† London,‡‡ and others.§§ According to them the quantum theory is rooted in the very fact that the measurement

* According to results kindly communicated to the author by letter.

† Herzberg, Ann. d. Phys., **85**, 565 (1927); cf. also L. Ornstein, Physikal. Zeitschr., **28**, 688 (1927).

‡ M. Born, Zeitschr. f. Phys., **38**, 803 (1926); **40**, 167 (1927).

§ Cf. the quotation in Jordan's paper, Zeitschr. f. Phys., **40**, 811 (1927).

|| E. Jordan, Zeitschr. f. Phys., **37**, 876 (1926); **40**, 809 (1927); **41**, 797 (1927); **44**, 1 (1927).

** W. Heisenberg, Zeitschr. f. Phys., **40**, 501 (1927); **43**, 172 (1927).

†† P. A. M. Dirac, Proc. Roy. Soc. (A), **113**, 621 (1927).

‡‡ F. London, Zeitschr. f. Phys., **40**, 198 (1926).

§§ M. Born, W. Heisenberg, P. Jordan, Zeitschr. f. Phys., **35**, 557 (1926).

of physical quantities is essentially characterised by a lack of precision, in that the accuracy of measurement of one of two canonically conjugate quantities is conditioned by that of the other (Heisenberg's Uncertainty Relation, *Ungenauigkeitsrelation*). In applying this idea to Schrödinger's theory the energy and the time co-ordinate are to be regarded as these two special quantities. In the proper states the energy is *sharply* determined; this makes the determination of the time impossible and the idea of orbits traversed in time falls out of the picture. Only the average time spent by the electron at every point, that is, the density of the charge-cloud, allows itself to be determined.

It is quite in conformity with the sense of these general theories that the ψ -function is itself a mathematical auxiliary quantity (*Hilfsgrösse*) and that it is only its *Norm* * (multiplied by e) that has a real physical meaning, namely the density of charge ρ . The ψ -function corresponds to what Jordan calls the "probability-amplitude" in contradistinction to the real probability, which is defined as the modulus of the probability-amplitude and is generally complex. Simple differential equations do not, however, hold for the moduli but for their linear factors, the amplitudes;—all this is analogous to the state of affairs in Schrödinger's theory.

We wish to go beyond this and although with less certainty point out a closer analogy, namely one which emerges from the relationships in the electromagnetic field: here the physical, that is, measurable quantities are the components of the energy-potential-tensor, that is, *quadratic* functions of the field intensities. Simple differential equations, namely those of Maxwell, do not, however, hold for the tensor components but for their linear factors, the intensities. These field intensities may also be regarded as mathematical auxiliary quantities introduced for calculating the actual physical relationships between the energy and the motion.

When we called the wave-mechanical density of charge a statistical mean of the electronic orbits previously used we did not imply that this mean is to be taken in the ordinary way. In wave-mechanics there is a finite density of charge, although only small, far outside the region of the earlier orbits, that is, in a region where ordinary statistics would lead to a zero value. Thus a new sort of statistics is involved which, indeed, is related to but not identical with the ordinary statistics of the orbits of classical mechanics.

B. Symmetry of the Charge-Distribution, in Particular in the Case of Complete Shells

When dealing with the Kepler motion, § 7, C, we emphasised that the S terms, that is, the states with $l = 0$, behave as if endowed with spherical symmetry, but the P and D terms ($l > 0$) behave as if they are only axially symmetrical. This symmetry not only simplifies the applications of the theory but, indeed, plays the decisive part in representing the physical facts (cf. the examples treated under E and F).

* See p. 34.

What holds for the hydrogen atom may be applied qualitatively to the more complicated atoms. In them, too, the S terms are *spherically symmetrical*. For if, following *Atombau*, Chap. VII, § 4, we imagine the dissimilarity of such atoms to hydrogen as produced by a centrally-symmetrical but non-Coulombian supplementary field, which represents the action of the atomic core on the series electron, this field can disturb the course of the qualitative values of the proper functions but not their central or axial symmetry. In other words, the inner central field will affect the radial component of the proper function but not the angular component $P_l^m(\cos \theta) e^{im\phi}$.

The spherical symmetry of the S-terms is immediately obvious in the case of the alkalis, where there is only one valency electron in a state for which $l = 0$; the same is true of the alkaline earths with their two valency electrons in the same state and quite generally for any S term due to the superposition of only spherically-symmetrical states ($l = 0$). The fact that a spherically-symmetrical state results from the superposition of states that belong individually to $l > 0$, so long as the resultant l is zero (namely, the group quantum number \bar{l} , see below, under D; $\bar{l} = 0$ is the general definition of the S term), cannot be proved here.

Further, according to A. Unsöld,* spherical symmetry is realised in the case of all *complete shells* ($j = 0$). What a complete shell is can be fully explained only by Pauli's† Exclusion Principle (*Verbot*), which was not known when the fourth edition of *Atombau* appeared, but which has since become of pre-eminent importance in all questions of atomic physics, in particular for the theory of the periodic system. It states that in the case in which several electrons occur *each well-defined state can be realised by at most one electron*. A state is said to be well defined when it is characterised by *four* quantum numbers. We best choose as this quartet of quantum numbers

$$n, l, m \text{ and } m_s.$$

The first three correspond to the three degrees of freedom of the electron in the hydrogen atom, such as we have hitherto taken into consideration, namely its three co-ordinates r, θ, ϕ ; and it is of no importance whether we choose the radial quantum number n_r for the r -degree of freedom or, instead, the more usual principal quantum number $n = n_r + l + 1$. But what does the fourth quantum number, which we have denoted by m_s , signify? What is its corresponding degree of freedom? The answer‡ was not directly given by Pauli's exclusion principle but its ground was prepared by it: this fourth degree of freedom lies in the electron itself, namely in the angular moment due to its spin, taken along the same axis as that to which the magnetic quantum number m (or expressed more fully m_l) is referred.

* Münchener Dissertation, 1927; Ann. d. Phys., **82**, 355 (1927).

† Zeitschr. f. Phys., **31**, 765 (1925).

‡ Hypothesis of Goudsmit and Uhlenbeck, Physica **5**, 266 (1925).

If we further add that the number m_s is to be capable* only of the two values $m_s = \pm \frac{1}{2}$, we may define the conception of a complete shell thus: it is formed by electrons of all those states m , m_s which for given values of n and l are possible according to Pauli's exclusion principle. The number of these electrons is $2(2l + 1)$, in which the factor 2 arises from the two possibilities for m_s , whereas the factor $2l + 1$ corresponds to the sum-total of integral values of m that result from the condition $-l \leq m \leq +l$.† We write down a first approximation to the proper function ψ and the density of charge ρ for any one of these electrons, that is we neglect the mutual action of the electrons among themselves (but assume the action of the inner central field to be taken up in the radial component R) and we also neglect a component which corresponds to the moment of momentum due to the spin:

$$\psi = \frac{R}{N_r} \frac{P_l^m(\cos \theta)}{N_\theta} \frac{e^{im\phi}}{N_\phi}, \quad \rho = \frac{R^2}{N_r^2 N_\theta^2} \frac{[P_l^m(\cos \theta)]^2}{N_\phi^2}$$

Here R and N_r are dependent on n and l but independent of m . We also assume that R does not appreciably depend on the fourth quantum number m_s . The same then holds of N_r , whereas N_ϕ is, as we know, equal to $\sqrt{2\pi}$, that is, it is constant. If we now sum up for the $2(2l + 1)$ electrons of the complete shell, we obtain, omitting all factors independent of m ,

$$\Sigma \rho = \frac{2R^2}{N_r^2 N_\theta^2} \sum_{-l}^{+l} \frac{1}{N_\phi^2} [P_l^m(\cos \theta)]^2 \quad (1)$$

The notation P_l^m here indicates that of course $P_l^{-m} = P_l^{+m}$. The density obtained by summation in this way is at the same time the density of charge that corresponds to the shell, to the same degree of approximation as that caused by neglecting the mutual action between the electrons. We assert that this density is independent of θ , that is, that it is spherically symmetrical.‡

The proof is based on a theorem of spherical harmonics which has long been known, namely, the *addition theorem* of spherical harmonics: if θ, ϕ and θ', ϕ' are two points on the unit sphere and Θ their spherical distance, such that

$$\cos \Theta = \cos \theta \cos \theta' + \sin \theta \sin \theta' \cos (\phi - \phi'),$$

* This is the hypothesis of Goudamit and Uhlenbeck; the magnetic moment of the electron becomes equal to a Bohr magneton. An explanation of the magnitude of the mechanical and magnetic moment of the electron will come quite naturally out of Dirac's theory (Chap. II, § 10).

† The use of negative values of m will be found very convenient in the sequel (cf. Note 1, p. 58).

‡ Unsöld (*loc. cit.*) starts from the theory of perturbations in proving the spherical symmetry. The simpler method of directly superposing the densities, used in the text, was indicated early by the author in the *Physik. Zeitschr.*, 23, 238 (1927).

then

$$P_l(\cos \Theta) = \sum_{-l}^{+l} \binom{l}{m} \frac{(l - |m|)!}{(l + |m|)!} P_l^{|m|}(\cos \theta) P_l^{|m|}(\cos \theta') e^{im(\phi - \phi')} \quad (2)$$

If in this expression we set $\theta = \theta'$ and $\phi = \phi'$ then $\cos \Theta$ becomes equal to 1 and we get

$$P_l(1) = \sum_{-l}^{+l} \binom{l}{m} \frac{(l - |m|)!}{(l + |m|)!} [P_l^{|m|}(\cos \theta)]^2 \quad (2a)$$

But by § 6, eqn. (30), the right side is identical with our sum in (1) except for the factor $\frac{2l+1}{2}$; by eqn. (12a) of § 2 the left side of (2a) is equal to 1. Hence we see that our distribution of charge (1) is in fact independent of θ , that is, spherically symmetrical.

The symmetrical form of (2) is generally distorted by writing

$$\left. \begin{aligned} P_l(\cos \Theta) &= P_l(\cos \theta) P_l(\cos \theta') \\ &+ 2 \sum_{-l}^{+l} \binom{l}{m} \frac{(l - m)!}{(l + m)!} P_l^m(\cos \theta) P_l^m(\cos \theta') \cos m(\phi - \phi') \end{aligned} \right\} \quad (3)$$

The advantage we gain by using negative m 's and the exponential forms becomes obvious by comparing (2) and (3). We can give our eqn. (2) an even more symmetrical form if we pass from the conventionally normalised P_l^m 's to the rationally normalised functions, that is, those divided by $N = N_\theta N_\phi$, which we shall call Π_l^m . (2) then shows itself to be equivalent to

$$\Pi_l(\cos \Theta) \Pi_l(1) = \sum_{-l}^{+l} \Pi_l^m(\cos \theta) \Pi_l^m(\cos \theta') e^{im(\phi - \phi')} \quad (4)$$

In this form the addition theorem comes out directly as a special case of a general theorem concerning a method of development in terms of proper functions, as A. Unsöld (*loc. cit.*) asserts, with reference to a lecture given by the author. Instead of this we may also prove eqn. (2) as follows. The left side of (2) is a spherical harmonic of l th degree symmetrical in θ, ϕ and θ', ϕ' , and continuous over the whole surface of the sphere, that is, it is a solution of the partial differential eqn. (1b) in § 2. The most general form of such a solution is given by the right side of (2), if we leave the numerical coefficient undetermined in each term: we call it c_m , and we must have $c_{-m} = c_m$. The c_m 's may now be successively determined thus. As in (2a) we set $\theta = \theta', \phi = \phi'$ and obtain, quite analogously,

$$\sum_{-l}^{+l} c_m [P_l^m(x)]^2 = 1, \quad x = \cos \theta \quad (5)$$

If we make $x = 1$, all P_l^m 's vanish for $|m| > 0$ on account of the factor $\sin^m \theta = (1 - x^2)^{m/2}$ and we get

$$c_0 [P_l(1)]^2 = c_0 = 1,$$

as it should be by eqn. (2) or (3). By taking the term with c_0 in (5) over to the right and dividing by $1 - x^2$, we get from eqn. (5) [cf. (25a) of § 6]

$$2c_1 [P_l'(x)]^2 + \dots = \frac{1 - [P_l(x)]^2}{1 - x^2} = P_l'(1) + \dots$$

in which the terms not written down vanish for $x = 1$ (the factor 2 is due to $m = \pm 1$). If we *again* make $x = 1$, it follows that

$$c_1 = \frac{1}{2P_l'(1)} = \frac{1}{l(l+1)},$$

cf. the expression (25) in § 6. But this agrees with the coefficient $\frac{(l-1)!}{(l+1)!}$ in (2). And so on. Eqn. (2) may therefore be regarded as proved in this way.

We now understand how much this spherical symmetry influences all considerations about the mutual action of complete shells. For a comparison we call to mind the model of the 8-shell developed by G. N. Lewis, namely the static octet. On account of its regular cubical arrangement this also has a high degree of symmetry such that the dipole and quadrupole moment of its charge-distribution vanishes. But there remain higher multipole moments, whose action some writers for a time wished to adduce to explain the forces in crystal lattices (cf. § 8, F). In opposition to this, however, our theorem states that for every kind of complete shells (stable molecules also in general possess complete shells) the multipole moments of any order vanish to a first approximation (that is, if we neglect the mutual action of the electrons).

C. The Incomplete Shells and their Magnetic Moments

We may regard the closed shells, like the former cube of electrons, as *electrostatic systems*: on account of their spherical symmetry there is no favoured direction for a possible flow of the charge. The case is different with the incomplete shells; like the former electronic orbits, these may be compared with systems of *stationary currents*. The reason for this is given by a relation which Schrödinger* interprets as the equation of continuity of electricity and uses for a general definition of the components of current.

Following Schrödinger we start out from the "time-equations" (13) and (13a) of § 5 which we write

$$\left. \begin{aligned} \frac{\partial u}{\partial t} &= \frac{h}{4\pi i m} \left(\Delta - \frac{8\pi^2 m}{h^2} V \right) u \\ \frac{\partial u^*}{\partial t} &= - \frac{h}{4\pi i m} \left(\Delta - \frac{8\pi^2 m}{h^2} V \right) u^* \end{aligned} \right\} \quad (6)$$

§ 8. The Incomplete Shells and their Magnetic Moment 89

We multiply the first equation by u^* , the second by u and add, observing that the member in V cancels,

$$\frac{\partial}{\partial t} u u^* = \frac{h}{4\pi i m} (u^* \Delta u - u \Delta u^*) \quad (6a)$$

To rewrite the right side in another form we recall the formal relation of Green's theorem: for any two functions u and u^* we have

$$u^* \frac{\partial^2 u}{\partial x^2} - u \frac{\partial^2 u^*}{\partial x^2} = \frac{\partial}{\partial x} \left(u^* \frac{\partial u}{\partial x} - u \frac{\partial u^*}{\partial x} \right)$$

$$u^* \Delta u - u \Delta u^* = \text{div} (u^* \text{grad } u - u \text{grad } u^*),$$

and so by (6a)

$$\frac{\partial}{\partial t} (u u^*) + \frac{h}{4\pi i m} \text{div} (u \text{grad } u^* - u^* \text{grad } u) = 0 \quad (7)$$

This relation has the form of the *hydrodynamic equation of continuity*. Just as we identified euu^* with the charge density ρ , we shall set

$$S = \frac{e}{m} \frac{h}{4\pi i} (u \text{grad } u^* - u^* \text{grad } u) \quad (8)$$

equal to the current density (corresponding to the hydrodynamic density of momentum $\rho \mathbf{v}$). Just as the equation of continuity asserts the conservation of mass in space and time, our e multiplied by eqn. (7) guarantees the *conservation of charge in space and time*: the charge ρ only alters according as it is brought up or conveyed away by the current. As Schrödinger emphasises, eqn. (7) at the same time guarantees the *persistence of the normalisation*. For if we integrate (7) throughout space, the second member vanishes; the second then asserts the independence of the quantity $\int u u^* d\tau$ from time; we were therefore justified in normalising it generally to unity.

The above reflections hold so far only for *one* electron. To be able to apply them to a system of several electrons, for example, to a shell of electrons, we recall the transition in § 1 from the wave-equation (11) for one electron to eqn. (12) for several particles. This transition can be applied to the above time-equations (6); the eqn. (7), which results from (6) remains preserved in form, but the symbol *div* now applies to the co-ordinates x_a, y_a, z_a of *all* particles (if the masses of the particles differ, the factor m must obviously be distinguished by an index a and must be written after the symbol *div*). Thus the conservation of charge so far occurs in multi-dimensional phase-space; S is similarly defined multi-dimensionally, in accordance with the generalised meaning of the "gradient." We can at once, however, arrive at a three-dimensional equation of conservation for the charge of a single particle if we integrate with respect to the co-ordinates of all the others. All members of the

expression under *div* that refer to the co-ordinates of these other particles vanish; the eqns. (7) and (8) and the signs *div*, *grad* again hold in the original three-dimensional sense if we take uu^* and $u \text{ grad } u^*$ and so forth to stand for the quantities integrated in the manner indicated. In this way we explain the rule already given earlier (§ 7, D) for determining the distribution of charge to be attributed to the individual particle and extend it by simultaneously determining the corresponding current flow S .

Just as in § 5, besides speaking of the density ρ of one state we also spoke of the density ρ_{nm} of the transition between the two states n and m , so here besides defining the current S that belongs to one state u , we also define a current S_{nm} which is calculated from two states n and m and hence belongs to the transition $m \rightarrow n$. For this purpose we attach to u in the first eqn. (6) the index n , and to u^* in the second equation the index m . We multiply the first equation by u_m^* , the second by u_n and get instead of (6a)

$$\frac{\partial}{\partial t}(u_n u_m^*) = \frac{\hbar}{4\pi i m} (u_m^* \Delta u_n - u_n \Delta u_m^*),$$

and hence, by Green's theorem, in place of (7)

$$\frac{\partial}{\partial t}(u_n u_m^*) + \frac{\hbar}{4\pi i m} \text{div} (u_n \text{grad } u_m^* - u_m^* \text{grad } u_n) = 0.$$

This signifies an equation of continuity for the density

$$\rho_{nm} = e u_n u_m^*.$$

Consequently we are justified in regarding the quantity

$$S_{nm} = \frac{e}{m} \frac{\hbar}{4\pi i} (u_n \text{grad } u_m^* - u_m^* \text{grad } u_n) \quad (8a)$$

as a generalised "current-density."

The quantities $S = S_{nn}$ defined in (8) form the diagonal elements of the "current matrix" S_{nm} , that is, a two-dimensional table of quantities, whose individual elements are allocated to the transitions (and those of the diagonal to the states themselves).

The current S as also S_{nm} will also have to be regarded as statistical quantities: S denotes the probable value of the charge which in unit time flows in the direction of the vector S through a unit surface placed at right angles to this direction.

We first apply eqn. (8) to the Kepler problem of one electron in a stationary state. For good reason, to be given later, we keep to our practice of writing the ψ -function in the complex form with the factor $e^{im\phi}$. For we see from the general form of the expression (8) that the component of S for every co-ordinate, on which ψ depends in a *real* way, must vanish. For S is, but for the factor $\frac{e}{m} \frac{\hbar}{2\pi}$, the *imaginary* part of $\psi \text{ grad } \psi^*$, and the latter is zero unless the differentiation is effected with

§ 8. The Incomplete Shells and their Magnetic Moment 91

respect to a co-ordinate which occurs as an imaginary in ψ . Let us take, for example, the co-ordinates r and θ in the case of the Kepler problem. Here we have

$$u = \psi \cdot e^{\frac{2\pi i}{h} Et}, \quad \psi = R \cdot P_l^m(\cos \theta) e^{im\phi}, \quad u^* = \psi^* e^{-\frac{2\pi i}{h} Et}$$

$$\psi^* = R \cdot P_l^m(\cos \theta) e^{-im\phi}, \quad \text{grad}_r = \frac{\partial}{\partial r}, \quad \text{grad}_\theta = \frac{1}{r} \frac{\partial}{\partial \theta}.$$

Thus,

$$u \frac{\partial u^*}{\partial r} = R \frac{dR}{dr} [P_l^m(\cos \theta)]^2$$

$$u \frac{\partial u^*}{r \partial \theta} = \frac{R^2}{r} P_l^m(\cos \theta) \frac{d}{d\theta} P_l^m(\cos \theta),$$

both of which expressions are real. Hence it follows from (8) that

$$S_r = S_\theta = 0.$$

The state of affairs is different, however, for the ϕ -direction. For this we have

$$\text{grad}_\phi = \frac{1}{r \sin \theta} \frac{\partial}{\partial \phi}$$

$$u \text{ grad}_\phi u^* = \frac{R^2 [P_l^m(\cos \theta)]^2}{r \sin \theta} (-im) = -\frac{im|\psi|^2}{r \sin \theta}$$

and by (8)

$$S_\phi = -\frac{e}{\mu} \frac{h}{2\pi} \frac{m|\psi|^2}{r \sin \theta} \quad (9)$$

in which we write μ for the electronic mass in order not to conflict with the quantum number m .

Hence, unless m happens to be equal to zero, the result is a stationary circular current about the polar axis $\theta = 0$, wherever the charge $|\psi|^2$ occurs, that is, really throughout infinite space.

From the current we pass on to its magnetic moment M , whose axis is of course the direction $\theta = 0$. We calculate it according to the rule: magnetic moment = current strength (in e.m.u.) times the surface enclosed by the current. Corresponding to the specific current S_ϕ is the current strength $S_\phi d\sigma$ per "cross-section" $d\sigma$ (normal to ϕ), or $\frac{1}{c} S_\phi d\sigma$ in e.m.u. The surface enclosed by the current is $\pi r^2 \sin^2 \theta$. The contribution of $d\sigma$ to the magnetic moment thus becomes

$$dM = -\frac{e}{\mu} \frac{h}{2\pi c} m|\psi|^2 \pi r \sin \theta d\sigma = -\frac{e}{\mu} \frac{h}{4\pi c} m|\psi|^2 d\tau,$$

where $d\tau = 2\pi r \sin \theta d\sigma$ signifies the volume of the circular tube formed by rotating $d\sigma$ about the axis $\theta = 0$. Summing up all such contributions we get, on account of the normalising condition $\int |\psi|^2 d\tau = 1$,

$$M = -\frac{e}{\mu} \frac{h}{4\pi c} m \quad (10)$$

Thus we get exactly m "Bohr magnetons" [cf. I, eqn. 14 p. 249; the negative sign signalises the negative charge of the electron]. We agree with E. Fermi* in regarding this simple wave-mechanical explanation of a relationship demanded by the earlier quantum theory as a beautiful confirmation of Schrödinger's assumption (8).

Before we discuss the general consequences of our result we must dispel a doubt that at once suggests itself. We obtained a current S differing from zero, as was emphasised, only from our form of ψ which was *complex* in ϕ . But there is nothing to prevent our passing on to the real part of ψ and to write $\cos m\phi$ or, more generally, $\cos (m\phi - \alpha)$ in place of $e^{im\phi}$. According to the above working we should then get $S_\phi = 0$. This apparent difficulty is to be explained as follows. Distinguishing our polar axis mathematically is only justified if this axis is also distinguished physically, for example, by a magnetic field applied in this direction. But the use of $e^{im\phi}$ is then no longer arbitrary but *necessary*. For in the Zeeman effect the magnetic field differentiates the two directions of revolution of the electron (that is, the two assumptions $e^{\pm im\phi}$) and procures for them different energy levels (proper values). Thus the above calculation is quite legitimate in the case of a magnetic field H and hence also in the case of passing to the limit $H \rightarrow 0$ (cf. I, Chap. IV, § 7, p. 242), in which the proper functions of the Zeeman effect merge into those of the Kepler problem, the direction of the magnetic field, however, remaining physically distinguished. The fact that wave-mechanics gives no electric current and no magnetic moment if we consider the degenerate Kepler problem, that is, if we do not consider it as a limiting case of a Zeeman problem and if we do not use the complex form then necessary of the proper function is no fault but rather a merit of wave-mechanics. Just as in the older theory the orbits could have any position in space in the absence of a magnetic field, so also the electric current of wave-mechanics has no definite orientation when no magnetic field is present. Since wave-mechanics always takes the statistical mean it can easily be understood that in the degenerate Kepler problem it arrives at the statement

$$S = 0, M = 0.$$

We can now answer the questions relating to current and magnetism in a complete or an incomplete shell. If we calculate to a first approximation, that is, neglecting the interaction between the electrons constituting the shell, then we may superpose the currents of these electrons just as under B we superposed their densities. We shall in fact see in the next section that the current and the density together form a higher entity (a "four-vector"). We also infer directly from (7) and the remarks that followed about multi-electronic systems that the superposition of currents occurs simultaneously with that of densities. But the magnetic moment is calculated from the currents. Consequently the magnetic moment is

* Nature, December, 1926, p. 876.

also obtained by superposing the magnetic moments of the individual electronic states. Thus it follows from (10) that

$$M = \Sigma m \text{ magnetons} \quad (11)$$

Hence an incomplete shell is in general paramagnetic; (11) gives the component of the paramagnetic moment in the direction of the magnetic field. We see at the same time that the complete shell is non-magnetic (or, better, diamagnetic); for in this case, in virtue of Pauli's exclusion principle, Σm becomes equal to zero, since in such a shell m assumes in turn the $2l + 1$ integral values between $-l$ and $+l$. The last assertion still remains unaffected if besides the magnetism of the currents S (of the "revolving electrons") we also include in our calculations the magnetism due to the rotation of the electrons themselves. For the latter contributions cancel in pairs since by Pauli's principle the total number of electrons in a complete shell is $2(2l + 1)$ and for every electron with m and $m_s = +\frac{1}{2}$ there is another with m and $m_s = -\frac{1}{2}$.

D. Individual l and Group Quantum Number \bar{l} . Spatial Quantising

The relationships described above may be simply epitomised as follows. In the case of one electron with a given l we interpret the l as a vector (moment of momentum of the revolution of the electron in the old sense or of the current in the present sense). The magnetic moment of the current in the direction of the z -axis is determined by the projection m of l on this axis. To the $2l + 1$ possible values of m , namely $-l \leq m \leq +l$ there correspond $2l + 1$ possible positions of l , as represented in Fig. 8, in which, of course, every position can be rotated about the z -axis on a cone. In this way we arrive at the analogy in wave-mechanics to the spatial quantising of Chap. II, § 8 (Vol. I). The flow is, so to speak, latent until we direct it by means of an applied magnetic field and even then shows only the component of its moment in the direction of the axis of the magnetic field, that is, the projection m of the moment vector on this axis. Our figure goes a little, but not essentially, beyond the analytical facts of wave-mechanics and summarises it in a convenient form.

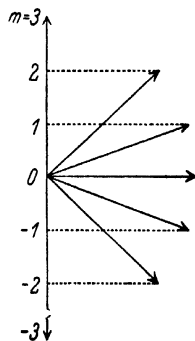


FIG. 8.

Spatial quantising. Magnetic quantum number m and azimuthal quantum number l for $l=3, m=0, \pm 1, \pm 2, \pm 3$.

Besides l we shall now take into account the electron spin s , that is, in addition to m_l (which we have hitherto denoted simply by m) we consider the magnetic quantum number m_s which $= \pm \frac{1}{2}$. The sum of these two, $m = m_l + m_s$, is then no longer integral with the maximum values $\pm l$, but half-integral with the maximum values $\pm (l + \frac{1}{2})$ and $\pm (l - \frac{1}{2})$. The number of values of m is no longer $2l + 1$ but $2(2l + 1)$ on account

of the additional possibility introduced by m_s having the two values $\pm \frac{1}{2}$, as we have already seen at the conclusion of the previous section C. We divide this number into two groups of $(2l + 2)$ and $(2l)$ values, respectively, and assign them to the two values

$$j = l + s = l + \frac{1}{2} \text{ and } j = |l - s| = |l - \frac{1}{2}| \quad . \quad (12)$$

respectively.

Thus by taking into account the spin of the electron we arrive by means of wave-mechanics (or pseudo wave-mechanics) at the "inner quantum number" j which played the principal part in systematising spectra in Chap. VIII, Vol. I. The new feature is that we now introduce this number for the individual electron itself, so that it occurs even for the H-atom, whereas we earlier believed it to be necessary only in dealing with the structure of multiplets.

If we imagine Fig. 8 to be constructed with the half-integral numbers $j = l \pm \frac{1}{2}$ in place of the whole number l , then the corresponding $m = m_l + m_s$ follows j in becoming half-integral and there are $2j + 1$ of them for each of the two values of j , that is, as above indicated, there are $2l + 2$ or $2l$ of them respectively. In the case $l = 0$ (S-term) the negative value of the two j -values, $j = l \pm \frac{1}{2}$, of course drops out, as is indicated in (12) by the vertical lines. Thus the S-term shows itself to be simple and all other terms double, like the alkali terms.

After Landé had repeatedly called attention to the alkali-like character of the Röntgen terms and after the Röntgen terms (since 1916) had been recognised as being hydrogen-like, the alkali-like doublet character of hydrogen terms was generally recognised in 1926 simultaneously with the discovery of the spinning electron. This necessitated numbering the fine-structure levels with two quantum numbers j (half-integral) and l (integral) instead of, as earlier, with the one quantum number $k (= l + 1)$. As a result of this the selection rules become modified and we get certain weak but experimentally proved components of the fine structure, as well as of the anomalous Zeeman effect and of the Paschen-Back effect—all in agreement with what occurs in the case of the alkalies. We shall revert to these points in the next section. Here we wish to emphasise only that the systematic structure of spectra becomes beautifully completed by the new classification of the hydrogen lines: the "Alternation Law" * (Wechselsatz, *Atombau*, Chap. VIII, § 2, and I, Chap. VI, p. 380) now holds without exception down to the first elements of the periodic system; lithium, which has a doublet spectrum, is followed by helium with singlet and triplet spectra and finally by hydrogen with a typical doublet spectrum.

We now pass on from the individual electron to a *group of electrons*, but for the present we shall leave out of account the electronic spin. The electrons may have the same l or different l 's and may belong to the same shell or to different shells (i.e. have the same or different n 's).

* *Three Lectures on Atomic Physics*, A. Sommerfeld, p. 89 (Methuen & Co., Ltd.)

They will be referred to the same magnetic axis: their mutual action (so far as it is not included in the central field mentioned on p. 85) will be neglected as in § B and C of this section. The currents S then simply add up arithmetically, and hence also the magnetic quantum numbers m [see eqn. (11) in C]. Thus we obtain for the "magnetic quantum number \bar{m} of the group"

$$\bar{m} = \sum m_\nu \quad . \quad . \quad . \quad . \quad . \quad (13)$$

where the index ν distinguishes the different electrons. If we now define the "azimuthal quantum number \bar{l} of the group" as the positive maximum of the possible values of \bar{m} (this is the same definition as, according to Fig. 8, holds for the individual l_ν in relation to the possible values of m_ν), so we have for \bar{l} the relation

$$\bar{l} = \sum \vec{l}_\nu \quad . \quad . \quad . \quad . \quad . \quad (14)$$

The arrow signifies that we must compound the l_ν 's *vectorially*, with the limitation that l_ν is to have only such positions with respect to the magnetic axis as make its projection on this axis integral ($= m_\nu$), in accordance with Fig. 8. Then we also get for the projection of \bar{l} on this axis an integral value, namely the value of \bar{m} from eqn. (13). Accordingly eqn. (14) is to express that \bar{l} can assume all positive *integral* values that may be formed by addition or subtraction of the l_ν 's:

$$\bar{l}_{\max} \geq \bar{l} \geq \bar{l}_{\min} \quad . \quad . \quad . \quad . \quad . \quad (14a)$$

As an example, we consider two electrons with, say,

$$l_1 = 1, \quad l_2 = 2.$$

Then, by (14a)

$$l_2 + l_1 \geq \bar{l} \geq |l_2 - l_1|,$$

thus

$$\bar{l} = 3 \text{ or } 2 \text{ or } 1.$$

In spectroscopic nomenclature this means either an F-term ($\bar{l} = 3$), a D-term ($\bar{l} = 2$), or a P-term ($\bar{l} = 1$). On account of the altered notation for the azimuthal quantum number $\bar{l} = k - 1$ we now get the following Table 2 in place of the earlier Table 47 (*Atombau*, p. 496), using capitals instead of small letters for the term-symbols (following an American suggestion):

TABLE 2.

$\bar{l} =$	0 S-	1 P-	2 D-	3 F-	4 G-	5... H-... Term
-------------	---------	---------	---------	---------	---------	--------------------

This table states that in the case of several electrons it is not the azimuthal quantum number l of a single "radiating electron" ("Leucht-elektron"), but the azimuthal group quantum number \bar{l} that is to determine the symbol and character of the term. This is logical since, for

example, the number of magnetic separation terms and hence also the quantum weight of the term (cf. *Atombau*, Chap. VIII, p. 651), the electron spin being disregarded for the present, are given by the group quantum number l . As a test we apply this to our example. The F-term has seven magnetic levels ($\bar{m} = \pm 3, \pm 2, \pm 1, 0$), the D-term five ($\bar{m} = \pm 2, \pm 1, 0$), the P-term three ($\bar{m} = \pm 1, 0$); in all there are fifteen levels. The same number results from combining all possible values of m_1 (namely three) with all possible values of m_2 (namely five).

The preceding remarks are, of course, very incomplete. The arrangement of terms into a system can be described more exactly on the basis of the more accurate theory of the magnetic electron, which we shall develop in Chap. II, § 10. We shall then find that the "orbital moment of momentum" (*Bahnimpuls*) of the individual electron is not equal to the azimuthal quantum number l , that is, a constant, but that only the quantity which appears in the place of j occurs as the real constant of integration. In the case of several electrons the same would be true of the group quantum number l and the inner quantum number j which is related to it. Nevertheless the preceding rules give the lines along which the theory of spectra has developed and according to which we shall have to be guided in future in dealing with complicated spectra.

E. Sphere of Action in Collisions. Supposed Double Refraction in the Magnetic Field

As we have seen under B the charge ρ is spherically distributed in the case of S-terms ($l = 0$) and in the case of a complete shell ($j = 0$). The first case occurs in the ground state of the hydrogen atom and of the analogous alkalis. The second case occurs, for example, in the ground state of helium and mercury but also in the case of the molecules such as H_2 , O_2 . Like the charge so also the fields of force of such atoms or molecules are spherically symmetrical. We must therefore conclude that collisions between atoms and molecules, since they are governed by the reactions of the fields of force, take place as in the classical kinetic theory of gases, that is, in the manner of isotropic spheres.

In contradistinction to this the symmetry of earlier atomic models was not the symmetry of a sphere but that of a disc, for example, in the ground state of the hydrogen atom, where the dimensions of the atom in the plane of the revolving electron would at any rate have to be greater than in the perpendicular direction.

In the Stern-Gerlach experiment we have a means of directing these supposed atom-discs parallel to each other. For if the paramagnetic moment (equal to one magneton in the case of the H-atom and of the alkalis in the ground state) were due to the revolving electron (in reality, it is due to the structure of the electron), the magnetic axis would have to be normal to the orbital plane and would have to set itself parallel or anti-parallel to the lines of force of the external magnetic field. This

adjustment has actually been proved to occur not only in the case of Ag and the alkalis (cf. *Atombau*, Chap. II, p. 146) but also in the case of the H-atom.* This phenomenon is also to be expected from the standpoint of wave-mechanics, as we saw in the preceding section, except that here it does not require the radius of action to be different in different directions, as on the older view.

An experiment by R. Fraser† with the positive rays of hydrogen has a bearing on this. The H^+ -particles (protons) are surrounded by a spherically symmetrical field, according to the older theory, as well as according to wave-mechanics. Their presence would therefore weaken the chance of a decision between the truth of these theories, so an electric field was used to deflect them from entering into the actual region of experiment. The neutral H-particles fell on to a thermopile and were counted by means of the deflection produced in an attached galvanometer; this was done with and without a magnetic field alternately, the magnetic lines of force being parallel to the direction of the positive ray. According to the older theory the collisions while the magnetic field was acting should occur more often and hence the galvanometer deflections should be less than when the field was off, as with the field on all H-atoms would present the whole face of their discs to the residual gas (H_2 or argon), whereas with the field off only a fraction of the disc, depending on the accidental orientation of the atom, would come into action as a cross-section effective in collision ("*Stossquerschnitt*"). Experimentally, however, no difference showed itself in the deflections when the field was on or off. This agrees with the new theory, according to which the H-atom is spherically symmetrical and hence its effective cross-section in collisions is not influenced by the magnetic orientation.

Even in his first fundamental notice about the theory of "quantising of direction in a magnetic field" ("*Richtungsquantelung im Magnetfelde*") Stern‡ concluded on the basis of the atomic models then in favour, that the magnetic orientation of atoms would necessarily lead to double refraction which, if it existed, should already have manifested itself in numerous earlier experiments, for example, on sodium vapour. This double refraction would be expected to be independent of the strength of the magnetic field and to occur also at distances far removed from the places characterised by anomalous dispersion. In fact, if all atoms, for example, in Na-vapour are magnetically directed, and if a light-ray is transmitted perpendicularly to the magnetic lines of force, then on the view that the atomic symmetry is that of a disc the two electric components into which the natural vibration of light can be resolved would lie differently with respect to the atom; the component parallel to the magnetic lines of force lies *normally* to the orbital plane, the other lies *in* the orbital plane. They would evoke different reactions in the valency

* E. Wreda, *Zeitschr. f. Phys.*, **41**, 569 (1927).

† *Proc. Roy. Soc.*, **114**, 212 (1927).

‡ *Zeitschr. f. Phys.*, **7**, 249 (1921).

electron and hence would have different velocities of propagation. In other words, magnetic double refraction should occur, independent of the magnetic field strength. It has been proved by various experimenters that this is not so. The last and most accurate experiments were carried out by R. Fraser* and W. Schütz† with a completely negative result. We already know how this is to be interpreted: this negative result does not deny the adjustment of atoms in a magnetic field, which has, indeed, been proved by Gerlach and Stern's experiment, but it is evidence against the view that atoms have a disc-like symmetry; in the spherically symmetrical structure of the Na-atom in the S-state this contradiction does not occur.

F. The Forces in Crystal Lattices

If we imagine a polar crystal, for example, Na, Cl, to be built up of ions of both signs, then *Coulombian attractions* are exerted between

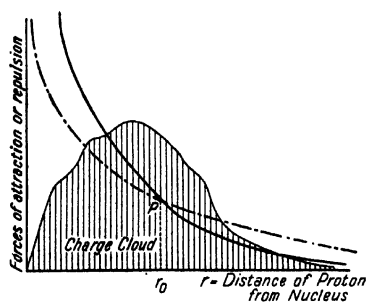


FIG. 9.—Diagrammatic explanation of the repulsive forces in the crystal lattice.

neighbouring atoms. The crystal would collapse into itself, if there were not also *repulsions* to balance these attractions. It was formerly believed that such repulsions could be derived from the cube conception (*Atombau*, Chap. III, p. 187). But this derivation was unsatisfactory as it assumed that the cubes were placed parallel to each other. Moreover, the multipole moments used in it do not

exist at all, according to wave-mechanics (cf. the end of § 8 B), since all the ions that come into question here consist of complete shells and have therefore spherical symmetry. The cube idea must be dropped, *not* because it is static—so is the charge-cloud of the complete shells, according to wave-mechanics—but because it has too little symmetry.

On the other hand, the idea of the continuous charge-cloud accounts without artificial assumptions for the repulsive forces and their actual magnitude. But we shall not take as our example Na Cl, in which the cation Na^+ and the anion Cl^- are approximately equal in size (Na^+ is smaller than neon, because its electron shell is contracted owing to its having excess of one in the nuclear charge, whereas Cl^- is somewhat larger than argon, since its outer electronic shell is extended owing to its defect of nuclear charge compared with argon), but we shall assume the

* Phil. Mag., 1, 885 (1926).

† Zeitschr. f. Phys., 32, 858 (1928).

cation to be a point or at any rate very small compared with the anion ; so we shall think of, say, Li Cl (the K-shell of Li⁺ is smaller even than that of He).

In Fig. 9 we represent diagrammatically the extended charge-cloud of the anion as results, for example, in the case of Cl⁻ from the superposition of the charge-clouds of the K-, L-, M-electrons and we follow the cation (proton or in our case Li-ion) while it penetrates into this charge-cloud. Let the Cl-nucleus with its charge of + 17 elementary units be at the origin of co-ordinates ; the charge-cloud represents in all the charge - 18. When the proton is at a great distance from the Cl-nucleus the charge-cloud is very thin (the proton here experiences the Coulomb attraction $\frac{1}{r^2}$ which corresponds to the difference of charge 18 - 17 = 1.

In the figure this attraction is drawn as a broken curve, not only for great distances from the nucleus but also by extrapolation for small distances, where it represents only a part of the action of the forces.

For when r is diminished the parts of the charge-cloud that lie outside the sphere of radius r drop out of the calculation, since a spherical shell carrying a charge distributed with spherical symmetry exerts no Coulomb force on a point inside it. Hence the attraction is diminished. Instead of this we can also say that a repulsion becomes superposed on

the Coulomb attraction $\frac{1}{r^2}$; this repulsion is represented in the figure by the continuous curve. For great distances of r it is small, like the density of the charge-cloud, but it increases as r decreases ; in our case it increases to the amount $\frac{17}{r^2}$ in the immediate vicinity of the nucleus.

Thus the repulsion finally outweighs the attraction and there is a distance r_0 corresponding to the point of intersection P of the broken and the continuous curve, where the repulsion and attraction balance. This is the position of equilibrium of the proton or of the Li⁺-ion in the charge-cloud of the anion. More extended cations such as Na⁺ will also find a similar position of equilibrium.

A. Unsöld * shows that this meets the facts by using as an approximation in the vicinity of $r = r_0$ a power of r^{-n} . According to Born and Landé's cube conception n should equal 5 for an ion with an outer L-shell and for a point-like cation. The wave-mechanical explanation, however, gives a value for n that increases systematically with r_0 and in parts exceeds 5 and in other parts falls below 5. The values obtained by observation do the same, as can be seen from the crystal lattices or, according to R. Mecke, from band-spectra.

* Zeitschr. f. Phys., **43**, 568 (1927).

§ 2. Relativistic Generalisations. The Principle of Variation. Application to the Kepler Problem

Nowadays we must demand of every mathematical-physical method that it satisfy the principle of relativity. In our case we may regard this as referring particularly to the postulate of invariance with respect to Lorentz transformations, since gravitational forces play no part in atomic events. The general principle of relativity therefore becomes restricted in our cases practically to the special theory of relativity.

Further, in extending wave-mechanics relativistically we shall deal only with the one-electron problem, not for reasons of simplicity but chiefly because the relativistic formulation of the many-electrons problem still seems to offer unresolved difficulties.

We shall proceed by first formally generalising the form of the wave-equation so far used: later we shall give a method of derivation in the sense of the optical analogy of § 1.

It is particularly important that at this juncture we shall learn how to fit magnetic forces (more generally, forces without potential) into the wave-equation.

A. The Relativistic Form of the Wave-equation

Instead of the co-ordinates x, y, z, t hitherto used, we now take the space-time co-ordinates

$$x_1 = x, \quad x_2 = y, \quad x_3 = z, \quad x_4 = ict.$$

In them eqn. (13) § 5 becomes

$$\sum_1^3 (k) \frac{\partial^2 u}{\partial x_k^2} + \frac{4\pi m_0 c}{h} \frac{\partial u}{\partial x_4} - \frac{8\pi^2 m_0}{h^2} \nabla u = 0 \quad (1)$$

in which we have written m_0 (the rest mass of the electron) in place of the previous m .

To represent the electromagnetic field acting on the electron we use the "four-potential" (*Viererpotential*) Φ , whose components Φ_1, \dots, Φ_4 we shall define* in terms of the usual vector potential \mathbf{A} and the scalar potential ϕ :

$$\Phi_1 = \frac{\alpha}{e} \mathbf{A}_x, \quad \Phi_2 = \frac{\alpha}{e} \mathbf{A}_y, \quad \Phi_3 = \frac{\alpha}{e} \mathbf{A}_z, \quad \Phi_4 = i \frac{\alpha}{e} \phi \quad (2)$$

where α is the fine-structure constant of I, eqn. (9), p. 213, so that

$$\frac{\alpha}{e} = \frac{2\pi e}{h c}.$$

Between and the potential energy V included in (1) [ϕ is referred to the charge 1, V to the charge e] there is the relation

$$e\phi = V - E_0, \quad E_0 = m_0 c^2 \quad (3)$$

* The factor $\frac{\alpha}{e}$ here inserted will be found useful later.

§ 9. The Relativistic Form of the Wave-equation 101

In actual fact the value $V = E_0$ = proper energy is that corresponding to the value $\phi = 0$ (field-free electron). From (2) and (3) it follows that

$$V = E_0 - i \frac{e^2}{\alpha} \Phi_4. \quad (3a)$$

We also mention the well-known subsidiary condition

$$\text{Div } \Phi = \sum_1^4 (k) \frac{\partial \Phi_k}{\partial x_k} = 0 \quad (3b)$$

which has to be imposed on \mathbf{A} and ϕ , when they are introduced, in order to make their definition unambiguous (*eindeutig*).

If we substitute (3a) in (1) and use the abbreviation

$$A = \frac{\alpha E_0}{e^2} = \frac{2\pi}{h} m_0 c \quad (4)$$

we get

$$\sum_1^3 \frac{\partial^2 u}{\partial x_k^2} + 2A \left(\frac{\partial u}{\partial x_4} + i \Phi_4 u \right) - 2A^2 u = 0 \quad (5)$$

We simplify the following calculations by introducing the differentiation symbol

$$\Omega_k = \frac{\partial}{\partial x_k} + i \Phi_k \quad (5a)$$

suggested by the factor of $2A$.

This symbol is to be used not only for $k = 4$ but also for $k = 1, 2, 3$. If, as in the problems so far considered, $\mathbf{A} = 0$, Ω_k , of course, becomes identical with $\partial/\partial x_k$ with $k = 1, 2, 3$. Our eqn. (5) referred only to such problems. We remain in agreement with it if we re-write it in the form

$$\left\{ \sum_1^3 \Omega_k^2 + 2A(\Omega_4 - A) \right\} u = 0 \quad (6)$$

So far we have not gone beyond our initial equation (1), nor have we yet attained our goal, namely to achieve symmetry in the four co-ordinates, $x_1 \dots x_4$, as is demanded by the principle of relativity. To do this we must modify (6) by terms that vanish when $c \rightarrow \infty$, or, as we may also say, by terms that are small compared with E_0 . We first note that since u depends on the time in the manner expressed in eqn. (11), § 5, we have

$$\frac{\partial}{\partial x_4} = \frac{2\pi}{hc} E = \frac{\alpha E}{e^2} \sim \frac{\alpha E_0}{e^2} \quad (6a)$$

Since further $e\phi \ll E_0$ [cf. (3)], we also have

$$|\Phi_4| \ll \frac{\alpha E_0}{e^2} \text{ and, cf. (6a), } \Omega_4 \sim \frac{\alpha E_0}{e^2}, \text{ i.e. } \Omega_4 \sim A.$$

Hence we may also set $2A \sim \Omega_4 + A$.

If we substitute this in (6), then

$$\left\{ \sum_1^3 \Omega_k^2 + \Omega_4^2 - A^2 \right\} u = 0$$

or

$$\left\{ \sum_1^4 \Omega_k^2 - A^2 \right\} u = 0 \quad . \quad . \quad . \quad (7)$$

The symmetry required by relativity has now been fully attained; what we have added to the original form of the wave-equation are only such correction terms as vanish in passing to the limit $c \rightarrow \infty$.

We could carry the symmetry of eqn. (7) a step further hypothetically by writing $A = i\Omega_0$ and taking Ω_0 as a symbol of differentiation with respect to a fifth co-ordinate x_0 . Then the five-dimensional wave-equation

$$\sum_0^4 \Omega_k^2 u = 0 \quad . \quad . \quad . \quad (7a)$$

would arise out of (7). A five-dimensional formulation of the theory of relativity was first suggested by Kaluza,* and carried a stage farther by Einstein† and O. Klein.‡

We return to the four-dimensional form (7) and easily convince ourselves that the quantity Ωu behaves like a four-vector. The fact that this is true of the *second* component of Ωu , namely $i\Phi u$ [cf. (5a)] is clear, since Φ was defined as a four-vector and u is a scalar quantity. But the first component

$$\frac{\partial u}{\partial x}$$

also behaves like a four-vector, and is, indeed, like Φ , a vector contravariant to the co-ordinate vector. However, in the special theory of relativity it is not necessary to distinguish between covariant and contravariant vectors (between parallel and orthogonal projection). Since then Ωu transforms itself like a four-vector, $\Sigma \Omega_k^2 u$ is invariant and $= \Sigma \Omega_k'^2 u$, if we take Ω_k' as standing for an operator formed from any orthogonally transformed co-ordinates x_k' and the corresponding potential components Φ_k' .

We effect the operations prescribed in (7) and write

$$\begin{aligned} \Omega_k u &= \frac{\partial u}{\partial x_k} + i\Phi_k u \\ \Omega_k^2 u &= \left(\frac{\partial}{\partial x_k} + i\Phi_k \right) \left(\frac{\partial u}{\partial x_k} + i\Phi_k u \right) \\ &= \frac{\partial^2 u}{\partial x_k^2} + 2i\Phi_k \frac{\partial u}{\partial x_k} + i \frac{\partial \Phi_k}{\partial x_k} u - \Phi_k^2 u. \end{aligned}$$

* Th. Kaluza, Sitzungsber. d. Preuss. Akad., 1921, p. 966.

† A. Einstein, *ibid.*, 1927, pp. 23 and 26.

‡ O. Klein, *Zs. f. Phys.*, 37, 895, 1926.

§ 9. The Relativistic Form of the Wave-equation 103

In summing with respect to the k -values the member containing $\frac{\partial \Phi_k}{\partial x_k}$ drops out on account of the condition (3b) and we get, in view of (4),

$$\sum_1^4 \frac{\partial^2 u}{\partial x_k^2} + 2i \sum_1^4 \Phi_k \frac{\partial u}{\partial x_k} - \left(\sum_1^4 \Phi_k^2 + \frac{\alpha^2 E_0^2}{e^4} \right) u = 0 \quad (8)$$

We consider two special cases :

(a) *An electrostatic field*, that is, $\Phi_1 = \Phi_2 = \Phi_3 = 0$, $e^2 \Phi_4 = i\alpha(V - E_0) = i\alpha U$, where U is independent of t and is normalised in the usual way ($U = 0$ at ∞). By introducing ψ in place of u , which is possible here, by setting

$$u = \psi e^{\frac{2\pi i}{h} Et}$$

we get from (8) after a simple reduction

$$\Delta \psi + \frac{4\pi^2}{h^2 c^2} [(E - U)^2 - E_0^2] \psi = 0 \quad (9)$$

(b) *A magnetic field* ($\mathbf{A} \neq 0$) in addition to an arbitrary electric field which may or may not depend on the time. We shall consider a more special case than (8) by neglecting the relativity correction terms.

We have then to complete eqn. (1) by adding those terms of (8) which contain \mathbf{A} , i.e. the terms

$$\begin{aligned} & \frac{2i\alpha}{e} (\mathbf{A} \text{ grad } u) - \frac{\alpha^2}{e^2} \mathbf{A}^2 u \\ & = \frac{4\pi ie}{hc} (\mathbf{A} \text{ grad } u) - \frac{4\pi^2 e^2}{h^2 c^2} \mathbf{A}^2 u. \end{aligned}$$

We shall strike out the term with \mathbf{A}^2 since even when the strongest magnetic fields are used (as was done by Kapitza) the quadratic effects here allowed for are inappreciable. On the other hand, we shall have to retain the first term although it has c in the denominator and so apparently vanishes for $c \rightarrow \infty$, because e/c is the charge, measured in e.m.u., which now takes the place of the e previously measured in e.s.u. Thus we have to add this first term to eqn. (1) and, remembering that $V = E_0 + U$, we get

$$\left. \begin{aligned} \Delta u + \frac{4\pi i}{h} \left\{ \frac{e}{c} (\mathbf{A} \text{ grad } u) - m_0 \frac{\partial u}{\partial t} \right\} \\ - \frac{8\pi^2 m_0}{h^2} (E_0 + U) u = 0 \end{aligned} \right\} \quad (10)$$

We cannot here pass on to a spatial ψ -equation; rather, it is characteristic of the magnetic effect that it produces a relationship involving space and time (cf. the Larmor precession).

B. The Optical Analogy in the Relativistic Case

We revert to the beginning of § 1. There (3) was the classical equation of energy, (4) the Hamiltonian equation. In place of it we now write the relativistic form of the energy equation and the corresponding Hamiltonian differential equation. By I, note 16, p. 607, the energy equation is, if we use Cartesian space co-ordinates instead of plane polar co-ordinates and an arbitrary U instead of the potential energy $U = -\frac{eE}{r}$ of the Kepler problem :

$$1 + \frac{1}{c^2 m_0^2} (p_x^2 + p_y^2 + p_z^2) = \left(1 + \frac{W - U}{E_0}\right)^2 = \left(\frac{E - U}{E_0}\right)^2 \quad (11)$$

where $E = E_0 + W$ denotes the total energy. By I, eqn. (11), p. 607,

$$\Delta_1 S = c^2 m_0^2 \frac{(E - U)^2 - E_0^2}{E_0^3} = \frac{1}{c^2} [(E - U)^2 - E_0^2] \quad (12)$$

The arguments of § 1, which there led to (5) and (7) apply here without change. Thus the right-hand side of (12) can now be regarded as the "square of the refractive index." From eqn. (6) in § 1 we therefore get, taking into account the meaning of k_0 in eqn. (10), § 1,

$$\Delta\psi + \frac{4\pi^2}{h^2 c^2} [(E - U)^2 - E_0^2] \psi = 0 \quad (13)$$

This has brought us to our eqn. (9) again, the relativistic equation of energy for the case $\mathbf{A} = 0$. In the same way we could have derived the general eqn. (8) from the optical analogy, if we had started out from the general form of Hamilton's principle (*Atombau*, eqn. (68), p. 801), which takes into consideration the vector potential \mathbf{A} .

C. The Relativistic Four-current

We take the term four-current* as meaning the vectorial composition of the specific electric current and the specific electric charge, namely

$$\mathbf{S} = (S_x, S_y, S_z, ic\rho) \quad (14)$$

We calculate it from Green's theorem, as in § 8, C. To illustrate the cogency of this procedure it is advisable first to consider any *arbitrary* linear partial differential expression of the second order and the function *adjoint* to it (*adjungiert*).

Let $A, B, \dots F$ be arbitrary functions of, say, two independent variables x and y ; these functions are to be such that they can be

*This term, introduced by M. von Laue for the electrodynamic vector ρ_0^{∇} , $ic\rho$ having four components is better than the term four-density previously introduced by the author. The actual, dimensionally correct four-current is obtained by multiplying this expression by c , that is, it is given by $(\rho\mathbf{v}, ic\rho)$. Our above definition in eqn. (14) corresponds with this.

differentiated sufficiently often for our purpose. The most general linear differential expression for the dependent variable u is

$$L(u) = A \frac{\partial^2 u}{\partial x^2} + 2B \frac{\partial^2 u}{\partial x \partial y} + C \frac{\partial^2 u}{\partial y^2} + D \frac{\partial u}{\partial x} + E \frac{\partial u}{\partial y} + Fu \quad (15)$$

We multiply it by a second dependent variable v and transform the product, term by term, by replacing the differentiations of u by those of v . For example we get

$$\begin{aligned} vA \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x} \left(vA \frac{\partial u}{\partial x} \right) - \frac{\partial u}{\partial x} \frac{\partial A v}{\partial x} \\ &= \frac{\partial}{\partial x} \left(vA \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial x} \left(u \frac{\partial A v}{\partial x} \right) + u \frac{\partial^2 (A v)}{\partial x^2} \\ vB \frac{\partial^2 u}{\partial x \partial y} &= \frac{\partial}{\partial x} \left(vB \frac{\partial u}{\partial y} \right) - \frac{\partial u}{\partial y} \frac{\partial B v}{\partial x} \\ &= \frac{\partial}{\partial x} \left(vB \frac{\partial u}{\partial y} \right) - \frac{\partial}{\partial y} \left(u \frac{\partial B v}{\partial x} \right) + u \frac{\partial^2 (B v)}{\partial x \partial y} \end{aligned}$$

or also

$$\begin{aligned} vB \frac{\partial^2 u}{\partial x \partial y} &= \frac{\partial}{\partial y} \left(vB \frac{\partial u}{\partial x} \right) - \frac{\partial u}{\partial x} \frac{\partial B v}{\partial y} \\ &= \frac{\partial}{\partial y} \left(vB \frac{\partial u}{\partial x} \right) - \frac{\partial}{\partial x} \left(u \frac{\partial B v}{\partial y} \right) + u \frac{\partial^2 (B v)}{\partial x \partial y}, \text{ etc.} \end{aligned}$$

By summation we get

$$vL(u) - uM(v) = \frac{\partial S_x}{\partial x} + \frac{\partial S_y}{\partial y} = \text{div } S \quad (16)$$

with the abbreviations

$$M(v) = \frac{\partial^2 (A v)}{\partial x^2} + 2 \frac{\partial^2 (B v)}{\partial x \partial y} + \frac{\partial^2 (C v)}{\partial y^2} - \frac{\partial (D v)}{\partial x} - \frac{\partial (E v)}{\partial y} + F v \quad (17)$$

$$\left. \begin{aligned} S_x &= A \frac{\partial u}{\partial x} v - u \frac{\partial A v}{\partial x} + B \frac{\partial u}{\partial y} v - u \frac{\partial B v}{\partial y} + u D v \\ S_y &= C \frac{\partial u}{\partial y} v - u \frac{\partial C v}{\partial y} + B \frac{\partial u}{\partial x} v - u \frac{\partial B v}{\partial x} + u E v \end{aligned} \right\} \quad (18)$$

M is the differential expression *adjoint* to L . If $M = L$, the expression is called *self-adjoint* (*selbstadjungiert*). The conditions for this are, as may easily be seen by working out (17),

$$\left. \begin{aligned} \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} &= D, & \frac{\partial B}{\partial x} + \frac{\partial C}{\partial y} &= E \\ \frac{\partial^2 A}{\partial x^2} + 2 \frac{\partial^2 B}{\partial x \partial y} + \frac{\partial^2 C}{\partial y^2} &= \frac{\partial D}{\partial x} + \frac{\partial E}{\partial y} \end{aligned} \right\} \quad (19)$$

This case occurs, for example, in the potential expression $L(u) = \Delta u$, where $A = C = 1$, $B = D = E = F = 0$ and $S = v \text{ grad } u - u \text{ grad } v$.

By integrating (16) over an arbitrary region σ of the x - y -plane having the boundary s we get the generalised form of Green's theorem: *

$$\int \{vL(u) - uM(v)\}d\sigma = \int S_n ds \quad . \quad . \quad . \quad (20)$$

where S_n signifies the normal component (reckoned positive in the outward direction) of the vector S along the bounding curve s .

The above remarks were intended to show above all how intimately the vector S is connected with the differential expressions L and M .

If we now apply the same process to our four-dimensional case, we get by (5a)

$$v\Omega_k u = \frac{\partial}{\partial x_k}(uv) - u \frac{\partial v}{\partial x_k} + i\Phi_k uv,$$

that is,

$$v\Omega_k u = -u\Omega_k^* v + \frac{\partial}{\partial x_k}(uv) \quad . \quad . \quad . \quad (21)$$

Here Ω_k^* denotes $\frac{\partial}{\partial x_k} - i\Phi_k$, the symbol of differentiation conjugate \dagger to Ω_k . By substituting $\Omega_k u$ for u in (21) and, in the same equation, exchanging u for v and Ω_k for Ω_k^* we get

$$\left. \begin{aligned} v\Omega_k^2 u &= -\Omega_k u \Omega_k^* v + \frac{\partial}{\partial x_k}(v\Omega_k u) \\ u\Omega_k^{*2} v &= -\Omega_k^* v \Omega_k u + \frac{\partial}{\partial x_k}(u\Omega_k^* v) \end{aligned} \right\} \quad (21a)$$

From (21a) we get by subtraction

$$v\Omega_k^2 u - u\Omega_k^{*2} v = \frac{\partial}{\partial x_k}(v\Omega_k u - u\Omega_k^* v) \quad . \quad . \quad . \quad (21b)$$

We sum up over k from 1 to 4, add the term $A^2 uv$ with a positive and a negative sign to the left side and multiply by a constant γ , which we shall dispose of presently. We get

$$\gamma \{v \sum (\Omega_k^2 - A^2)u - u \sum (\Omega_k^{*2} - A^2)v\} = \text{Div } S \quad . \quad . \quad (22)$$

$$\begin{aligned} S_k &= \gamma \{v\Omega_k u - u\Omega_k^* v\} \\ &= \gamma \left\{ v \frac{\partial u}{\partial x_k} - u \frac{\partial v}{\partial x_k} + 2i\Phi_k uv \right\} \quad . \quad . \quad (22a) \end{aligned}$$

* For the application of this theorem to the general theory of problems in boundary values (unique solutions, representation in Green's functions), cf. Enzykl. d. Math. Wiss. II, A, 7c, article by Sommerfeld.

\dagger In the sense that x_k and Φ_k are to be regarded as real. For $k = 4$, only the form of Ω_4^* , not the numerical value, is conjugate to Ω_4 .

Eqn. (22) states : the expression $M(v)$ adjoint to our differential expression (7)

$$L(u) = \sum (\Omega_k^2 - A^2)u \quad . \quad . \quad . \quad (22b)$$

is the expression

$$M(v) = \sum (\Omega_k^{*2} - A^2)v = L^*(v) \quad . \quad . \quad . \quad (22c)$$

conjugate to L . If u satisfies the equation $L = 0$, then $v = u^*$ satisfies the adjoint equation $L^*(u^*) = 0$. At the same time the equation of continuity

$$\text{Div } S = 0 \text{ with } S_k = \gamma \left\{ u^* \frac{\partial u}{\partial x_k} - u \frac{\partial u^*}{\partial x_k} + 2i\Phi_k u u^* \right\} \quad . \quad (23)$$

holds.

We shall choose the constant γ so that if the relativity-corrections be neglected, and for a dependence on time of the form

$$u = \psi e^{\frac{2\pi i}{h} Et}, \quad u^* = \psi^* e^{-\frac{2\pi i}{h} Et} \quad . \quad . \quad . \quad (23a)$$

the fourth component of S becomes, in accordance with (14),

$$S_4 = ic\rho = ice\psi\psi^* \quad . \quad . \quad . \quad (24)$$

Now, on account of (23a), we have

$$u^* \frac{\partial u}{\partial x_4} - u \frac{\partial u^*}{\partial x_4} = \frac{4\pi}{ch} E\psi\psi^* \quad . \quad . \quad . \quad (23b)$$

and, on the other hand, on account of (3a)

$$2i\Phi_4 u u^* = \frac{4\pi}{ch} (E_0 - V)\psi\psi^* = -\frac{4\pi}{ch} U\psi\psi^* \quad . \quad . \quad (23c)$$

But $U \ll E_0$ and $E \sim E_0$. Consequently (23c) is to be neglected in comparison with (23b). If in (23b) we replace E by $E_0 = m_0 c^2$, which is allowed to the same degree of approximation, the postulate (24), in which we substitute for S_4 from (23) and (23b), yields

$$\gamma \frac{4\pi m_0 c}{h} \psi\psi^* = ice\psi\psi^*,$$

that is,

$$\gamma = \frac{e}{m_0} \frac{ih}{4\pi} \quad . \quad . \quad . \quad (24a)$$

Thus the final definition of the four-current is

$$S_k = \frac{e}{m_0} \frac{h}{4\pi i} \left\{ u \frac{\partial u^*}{\partial x_k} - u^* \frac{\partial u}{\partial x_k} - 2i\Phi_k u u^* \right\} \quad . \quad . \quad (25)$$

The first three components merge directly into the current S , eqn. (8) of § 8, defined earlier, if the term in $\Phi_k = \frac{A}{e}$ is neglected, as it obviously vanishes for $c \rightarrow \infty$ on account of the factor a . Through our choice of γ we have arranged that the fourth component becomes $ic\rho$ for $c \rightarrow \infty$. We may therefore say that our earlier definitions of current and density

are shown to be justified by our relativistic generalisation, as they are combined together into the one entity four-current and are intimately connected with the wave-equation.

D. The Principle of Variation

We start from the two eqns. (21a) and form from them

$$\begin{aligned} v \left\{ \sum \Omega_k^2 - A^2 \right\} u &= - \left\{ \sum \Omega_k u \Omega_k^* v + A^2 uv \right\} + \sum \frac{\partial}{\partial x_k} (v \Omega_k u) \\ u \left\{ \sum \Omega_k^2 - A^2 \right\} v &= - \left\{ \sum \Omega_k v \Omega_k^* u + A^2 uv \right\} + \sum \frac{\partial}{\partial x_k} (u \Omega_k^* v) \end{aligned} \quad (26)$$

In view of the notation in (22 b, c) we write for this

$$\begin{aligned} vL(u) + \Lambda &= \text{Div } P \\ uM(v) + \Lambda &= \text{Div } Q \end{aligned} \quad (26a)$$

Here

$$\Lambda = \sum \Omega_k u \Omega_k^* v + A^2 uv \quad (27)$$

a bilinear form in the quantities u , $\partial u / \partial x_k$ on the one hand and of the quantities v , $\partial v / \partial x_k$ on the other. P and Q are four-vectors

$$P = v \Omega u, \quad Q = u \Omega^* v \quad (27a)$$

which are related to the four-vector S in (22a) as follows

$$S = \gamma(P - Q) \quad (27b)$$

To recognise the general significance of these relationships, it is advisable once again to consider an arbitrary linear differential expression L , for example, in two variables x, y . That is, we start from eqn. (15) and obtain, after multiplication by v and transforming *once* (not, as above, twice) in the sense of "integration by parts,"

$$\begin{aligned} vL(u) + \Lambda &= \text{div } P \\ \Lambda &= \frac{\partial u}{\partial x} \left(\frac{\partial A v}{\partial x} + \frac{\partial B v}{\partial y} \right) + \frac{\partial u}{\partial y} \left(\frac{\partial B v}{\partial x} + \frac{\partial C v}{\partial y} \right) + u \left(\frac{\partial D v}{\partial x} + \frac{\partial E v}{\partial y} - F v \right) \\ P_x &= v \left(A \frac{\partial u}{\partial x} + B \frac{\partial u}{\partial y} + D u \right), \quad P_y = v \left(B \frac{\partial u}{\partial x} + C \frac{\partial u}{\partial y} + E u \right) \end{aligned} \quad (28)$$

In the same way, by starting from the adjoint expression $M(v)$, eqn. (17), we get (applying "integration by parts" in the case only to the first three summands)

$$\begin{aligned} uM(v) + \Lambda &= \text{div } Q \\ Q_x &= u \left(\frac{\partial (A v)}{\partial x} + \frac{\partial (B v)}{\partial y} \right), \quad Q_y = u \left(\frac{\partial (B v)}{\partial x} + \frac{\partial (C v)}{\partial y} \right) \end{aligned} \quad (28a)$$

Λ is here a bilinear expression in u, v and their derivatives, which is simultaneously associated with the two differential expressions L and M . If $M = L$, we have the self-adjoint case, and Λ becomes symmetrical in u and v . For, on account of (19), we can then write

$$\Lambda = \left. \begin{aligned} & A \frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + B \left(\frac{\partial u}{\partial x} \frac{\partial v}{\partial y} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial x} \right) + C \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \\ & + D \frac{\partial uv}{\partial x} + E \frac{\partial uv}{\partial y} - Fuv \end{aligned} \right\} \quad (29)$$

At the same time, as can also be shown from (19), $Q(u, v) = P(v, u)$.

Integrate (28) and (28a) over a closed region σ in the x - y -plane with the boundary s and the external normal n . By Gauss's theorem we have

$$\left. \begin{aligned} \int v L(u) d\sigma + \int \Lambda d\sigma &= \int P_n ds \\ \int u M(v) d\sigma + \int \Lambda d\sigma &= \int Q_n ds \end{aligned} \right\} \quad (30)$$

This pair of equations represents a second form of the generalised Green theorem and constitutes the counterpart to the first form contained in (20). If L is self-adjoint, this pair of equations reduces to *one* statement; for example, in the case of the equation of potential $A = C = 1$, $B = D = E = F = 0$, to the well-known equation

$$\int v \Delta u d\sigma + \int \left(\frac{\partial u}{\partial x} \frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \frac{\partial v}{\partial y} \right) d\sigma = \int v \frac{\partial u}{\partial n} ds,$$

or, respectively, the equation resulting from exchanging u and v in it.

We now form the variation of, say, the first of the two eqns. (30), varying u and v in the interior of the region of integration by arbitrary (continuous and sufficiently small) amounts δu , δv , but keeping u and v as also their first derivatives constant at the edge. We obtain

$$\delta \int \Lambda d\sigma = - \int \delta v L(u) d\sigma - \int v L(\delta u) d\sigma. \quad (31)$$

We transform the second integral on the right by means of eqn. (20). If in (20) we replace u by δu and take into account that S , according to its definition in eqn. (16), vanishes at the boundary on account of $\delta u = 0$, it follows from (20) that

$$\int v L(\delta u) d\sigma = \int \delta u M(v) d\sigma.$$

Substitution in (31) gives

$$\delta \int \Lambda d\sigma = - \int \delta v L(u) d\sigma - \int \delta u M(v) d\sigma. \quad (32)$$

The variation of the second eqn. (30) leads to the same result.

From (32) it follows: *the variation postulate*

$$\delta \int \Lambda d\sigma = 0 \quad (33)$$

entails, on account of the arbitrariness of δu and δv , the fulfilment of the two differential equations

$$L(u) = 0, \quad M(v) = 0 \quad (33a)$$

in the whole interior of the region of integration. The state of affairs resembles that of classical mechanics, where the Hamiltonian variation principle

$$\delta \int \Lambda dt = 0, \quad \Lambda = \text{Lagrange's function} = E_{\text{kin}} - E_{\text{pot}}$$

is equivalent to the differential equations of mechanics.

The application of this variation principle to wave-mechanics now presents itself at once. If the eqns. (26a) are integrated over a four-dimensional region $d\tau$ with the three-dimensional surface element $d\sigma$, we get the "second form of Green's theorem"

$$\left. \begin{aligned} \int vL(u)d\tau + \int \Lambda d\tau &= \int P_n d\sigma \\ \int uM(v)d\tau + \int \Lambda d\tau &= \int Q_n d\sigma \end{aligned} \right\} \quad (34)$$

The "first form of Green's theorem" which results from the integration of (22) (in which the factor γ may be set equal to 1), would run

$$\int vL(u)d\tau - \int uM(v)d\tau = \int S_n d\sigma \quad (34a)$$

Variation is then performed on the first (or the second) of eqns. (34) by varying u by δu and v by δv , but keeping fixed the limits u, v , of the region of integration and their first derivatives. This gives

$$\delta \int \Lambda d\tau = - \int \delta v L(u) d\tau - \int v L(\delta u) d\tau \quad (34b)$$

By eqn. (34a), however, we have, if u is replaced by δu , which causes the right-hand side to vanish

$$\int v L(\delta u) d\tau = \int \delta u M(v) d\tau.$$

Hence it follows from (34b) that

$$\delta \int \Lambda d\tau = - \int \delta v L(u) d\tau - \int \delta u M(v) d\tau \quad (35)$$

This equation contains the following *variation principle of wave-mechanics*. On account of the arbitrariness of $\delta u, \delta v$ the postulate

$$\delta \int \Lambda d\tau = 0 \quad (36)$$

is equivalent to the two differential equations

$$L(u) = 0, \quad M(v) = L^*(u^*) = 0.$$

Schrödinger* calls Λ the "Lagrange function of wave-mechanics."

As Gordon first pointed out (see the conclusion of this section), the

* *Collected Papers, The Energy-Momentum Theorem for Material Waves*, p. 180.

components of the four-current may be obtained by differentiating with respect to Φ . For, by (22a) and (27),

$$S_k = i\gamma \frac{\partial \Lambda}{\partial \Phi_k} \quad (37)$$

The advantage of the variation principle consists, as in ordinary mechanics, in the fact that it is independent of the special choice of co-ordinates. That is why it will be useful to us in § 11 in transforming the wave-equation into arbitrary curvilinear co-ordinates. The variation principle has also a practical value for numerical calculations; cf. the end of this section.

Schrödinger placed at the head of his first paper the more special variation principle

$$\delta \int \Lambda d\tau = 0, \quad \Lambda = \left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial y} \right)^2 + \left(\frac{\partial \psi}{\partial z} \right)^2 - \frac{8\pi^2 m}{h^2} (W - U) \psi^2 \quad (38)$$

in which $d\tau$ indicates integration over infinite *three*-dimensional space. We easily convince ourselves that this form of the variation principle comes out of (36) if we neglect the relativity corrections and pass to the simplest case for which $\mathbf{A} = 0$, and hence [cf. (3b)] U is independent of t .

Finally we can follow Schrödinger* and give the variation principle (38) the following more satisfactory form:

$$\delta \int H d\tau = 0, \quad H = \frac{1}{2m} \left(\frac{h}{2\pi} \right)^2 \left[\left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial y} \right)^2 + \left(\frac{\partial \psi}{\partial z} \right)^2 \right] + U \psi^2 \quad (38a)$$

with the normalising condition

$$\int \psi^2 d\tau = 1 \quad (38b)$$

The notation H is to call to mind the "Hamiltonian function" of ordinary mechanics, that is the sum of the kinetic and potential energies, the first regarded as a function of the *momenta* p :

$$H = T + U, \quad T = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2).$$

Our quantity H defined in (38a) arises from this, if we multiply each term by ψ^2 and if we replace

$$p_x \psi \text{ by } \frac{h}{2\pi} \frac{\partial \psi}{\partial x}, \text{ etc.}$$

(cf. also Chap. II, § 9).

By introducing a Lagrangian multiplier, called $-W$, say, we can combine (38a, b) in

$$\delta \int (H - W \psi^2) d\tau = 0 \quad (38c)$$

* *Collected Papers*, p. 2.

which is identical with (38) except for an irrelevant constant factor. The form (38a) of the variation principle is more satisfactory than (38), because the energy W does not occur in it but is determined as a multiplier by the process of variation itself.

Starting out from (38a, b) and using a method due to Ritz, G. Kellner* has calculated the numerical value of the ground term of He and Li^+ . E. Hylleraas† later obtained the ground term of helium by the same method but with increased accuracy; his result differs from the experimental values by only one hundredth per cent., a sign that the methods of wave-mechanics are capable of dealing even with the many-electron problem. If we wished to treat the same problem analytically (by the theory of perturbations) by starting from the differential equation, instead of numerically by means of the variation principle, we should have to remain satisfied with a much lower degree of accuracy.‡ We shall revert to this point in Chap. II, § 8.

E. The Relativistic Kepler Problem ✓

Since in this case the potential energy is independent of t , we may take as our basis the wave-equation in the simple form (9) with

$$U = - \frac{Ze^2}{r}.$$

The integration is performed exactly as in § 7. Assuming that

$$\psi = R P_l^m(\cos \theta) e^{im\phi}$$

we get for R the differential equation obtained earlier:

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left(A + 2\frac{B}{r} + \frac{C}{r^2} \right) R = 0 \quad (39)$$

but the constants A, B, C now have a different meaning from that in § 7. Their present significance follows from the equation of definition [cf. (9)]

$$A + 2\frac{B}{r} + \frac{C}{r^2} = \frac{4\pi^2}{h^2 c^2} \left[\left(E + \frac{Ze^2}{r} \right)^2 - E_0^2 \right] - \frac{l(l+1)}{r^2}$$

from which we infer

$$A = \frac{4\pi^2}{h^2 c^2} (E^2 - E_0^2), \quad B = \frac{4\pi^2}{h^2 c^2} Ze^2 E, \quad C = -l(l+1) + \frac{4\pi^2 Z^2 e^4}{h^2 c^2} \quad (40)$$

These are essentially the same values as in the older theory, I, eqns. (4a) p. 609, Note 16. We recognise this immediately if in (40) we set

$$E = E_0 + W, \quad E_0 = m_0 c^2 \quad (40a)$$

* Zeitschr. f. Phys., 44, 91 and 110 (1927).

† Ibid., 48, 469 (1928), and 54, 847 (1929).

‡ But compare J. C. Slater, Proc. Nat. Acad., 13, 428 (1927).

for then we get

$$\left. \begin{aligned} A &= \frac{4\pi^2}{h^2} \left(2m_0 W + \frac{W^2}{c^2} \right), \quad B = \frac{4\pi^2}{h^2} m_0 Z e^2 \left(1 + \frac{W}{m_0 c^2} \right), \\ C &= - \frac{4\pi^2}{h^2} \left(\frac{l(l+1)h^2}{4\pi^2} - \frac{Z^2 e^4}{c^2} \right). \end{aligned} \right\} \quad (40b)$$

These equations differ from the older equations just mentioned only in the unimportant addition of the common factor $4\pi^2/h^2$ and in having k^2 replaced by $l(l+1)$, a matter of greater importance. Our present expressions (40b), of course, merge into the values (3a) of § 7 when $c \rightarrow \infty$ (except in having m in place of m_0 and E instead of W , the latter corresponding to the normalisation of E and V there adopted).

To integrate (39) we proceed as in § 7 and substitute from eqns. (4), (4a), and (5)

$$A = -\frac{1}{r_0^2}, \quad \rho = \frac{r}{r_0}, \quad R = e^{-\rho^2/2} v. \quad (41)$$

and we get for v the differential equation

$$v'' + \left(\frac{2}{\rho} - 1 \right) v' + \left[\left(\frac{B}{\sqrt{-A}} - 1 \right) \frac{1}{\rho} + \frac{C}{\rho^2} \right] v = 0. \quad (42)$$

It differs from eqn. (7) of § 7 only in having instead of $-l(l+1)$ the quantity C , which differs from it by a correction member due to relativity. For by (40)

$$C = -l(l+1) + \alpha^2 Z^2,$$

where α denotes the fine structure constant. To solve (42) we make the assumption

$$v = \rho^\gamma \sum a_\nu \rho^\nu \quad (43)$$

as earlier in eqn. (7), § 7.

γ is here given by the characteristic equation [cf. (7b) of § 7]

$$\gamma(\gamma+1) = -C = l(l+1) - \alpha^2 Z^2$$

from which it follows that

$$\begin{aligned} \left(\gamma + \frac{1}{2} \right)^2 &= \left(l + \frac{1}{2} \right)^2 - \alpha^2 Z^2, \\ \gamma &= \sqrt{\left(l + \frac{1}{2} \right)^2 - \alpha^2 Z^2} - \frac{1}{2}. \end{aligned} \quad (44)$$

The decisive step is now to set up the recurrence formula for the coefficients a_ν in the expansion (43). Just as in eqn. (8), § 7, the factor of a_ν in the recurrence formula becomes

$$\nu + \gamma + 1 - \frac{B}{\sqrt{-A}}.$$

If we equate it to zero for $\nu = n_r$, the sum in (43) becomes a polynomial of degree n_r . Taking (44) into account we get

$$\frac{B}{\sqrt{-A}} = n_r + \sqrt{\left(l + \frac{1}{2} \right)^2 - \alpha^2 Z^2} + \frac{1}{2} \quad (45)$$

But by (40)

$$-\frac{B^2}{A} = -\frac{\alpha^2 Z^2}{E^2 - E_0^2} \quad (45a)$$

If we temporarily set

$$\beta = \frac{B}{\sqrt{-A}} \quad (45b)$$

we get by (45a)

$$\frac{E^2 - E_0^2}{E^2} = -\frac{\alpha^2 Z^2}{\beta^2}, \quad \frac{E_0^2}{E^2} = 1 + \frac{\alpha^2 Z^2}{\beta^2}.$$

If we raise the last equation to the $(-\frac{1}{2})$ th power and insert the value of β , we get [cf. also (40a)]

$$\frac{E}{E_0} = 1 + \frac{W}{m_0 c^2} = \left\{ 1 + \frac{\alpha^2 Z^2}{[n_r + \sqrt{(l + \frac{1}{2})^2 - \alpha^2 Z^2 + \frac{1}{4}}]^2} \right\}^{-1/2} \quad (46)$$

We recognise the general structure of the fine-structure formula, as previously derived in I, Chap. VIII, § 2, eqn. (23), p. 472, or I, note 16, p. 608. But there are appreciable differences: in place of the earlier $k^2 = (l + 1)^2$ we now have $(l + \frac{1}{2})^2$ under the root sign, and, in a certain sense, as a compensation for this the factor $\frac{1}{2}$ appears outside the root sign. The effect of this compensation is to make the *Balmer term* come out *correctly*, that is, as in § 7 without the relativity correction: for, by (46) we get as a first approximation, when $\alpha^2 = 0$,

$$W = -\frac{1}{2} \frac{m_0 c^2 \alpha^2 Z^2}{n^2} = -\frac{R h Z^2}{n^2}$$

with

$$R = \frac{1}{2} \frac{m_0 c^2 \alpha^2}{h} = \frac{2\pi^2 m_0 e^4}{h^3}.$$

But the *fine structure* is *wrongly* represented by (46). For, by (45), (45b), and (46) we have as a second approximation

$$\begin{aligned} \beta &= n - \frac{1}{2} \frac{\alpha^2 Z^2}{l + \frac{1}{2}}, \quad \frac{1}{\beta^2} = \frac{1}{n^2} \left(1 + \frac{\alpha^2 Z^2}{n(l + \frac{1}{2})} \right) \\ W &= -\frac{R h Z^2}{n^2} \left\{ 1 + \frac{\alpha^2 Z^2}{n^2} \left(\frac{n}{l + \frac{1}{2}} - \frac{3}{4} \right) \right\} \quad (47) \end{aligned}$$

If we calculate the frequency difference $-\Delta W/h$ of the fine-structure components for $l = 0, 1, \dots, n-1$ (n being kept fixed) we get

$$\begin{aligned} \text{for } n = 2, \quad & \frac{R \alpha^2 Z^4}{2^4} (4 - \frac{4}{3}) = \frac{8}{3} \Delta \nu \\ \text{for } n = 3, \quad & \begin{cases} \frac{R \alpha^2 Z^4}{3^4} (6 - 2) = 8 \Delta \nu_1 \\ \frac{R \alpha^2 Z^4}{3^4} (2 - \frac{6}{5}) = \frac{8}{15} \Delta \nu_2 \end{cases} \end{aligned}$$

$$\text{for } n = 4, \quad \begin{cases} \frac{R\alpha^2 Z^4}{4^4} (8 - \frac{8}{3}) = 16\Delta\nu_1 \\ \frac{R\alpha^2 Z^4}{4^4} (\frac{8}{3} - \frac{8}{5}) = \frac{8}{5}\Delta\nu_2 \\ \frac{R\alpha^2 Z^4}{4^4} (\frac{8}{5} - \frac{8}{7}) = \frac{8}{35}\Delta\nu_3 \end{cases}$$

The $\Delta\nu$, $\Delta\nu_1$, $\Delta\nu_2$, . . . on the right-hand side denote the differences of wave-number which in I, Chap. VIII, pp. 478, 479, we calculated according to the old formula; in particular, for $Z = 1$ $\Delta\nu$ is the well-known hydrogen doublet $\Delta\nu_H$, eqn. (1) of I, p. 481. Thus the factors in front of $\Delta\nu$, $\Delta\nu_1$. . . express by how much the new formula differs from the old theory. Since the old formula is fully confirmed by experiment the new formula proves to be *appreciably false*.

There is yet another objection* that can be urged against our calculation. The S-state ($l = 0$) exhibits a peculiarity which we must regard as depriving this state of its existence. For, by eqn. (44), we have for $l = 0$

$$\gamma = \sqrt{\frac{1}{4} - \alpha^2 Z^2} - \frac{1}{2} = \frac{1}{2}[(1 - 4\alpha^2 Z^2)^{\frac{1}{2}} - 1] = -\alpha^2 Z^2 + \dots < 0.$$

Hence by (43) v becomes ∞ for $\rho = 0$ (very slowly, on account of the exponent $\alpha^2 = 5 \cdot 10^{-5}$, but nevertheless ∞). Thus, according to our definition of proper functions, all states with $l = 0$ would, strictly speaking, be forbidden.

It is profitable to compare graphically the energy levels here found with the true levels. We set $Z = 1$ and on the left of Fig. 10 we indicate the position of the unresolved Balmer terms $\nu = R/n^2$ (without taking account of their absolute magnitude). Next to them we have drawn the correct fine structure according to *Atombau*, Chap. VI, p. 420, eqn. (6). We have on the one hand the "relativity correction for circular orbits" $k = n$, which displaces the corresponding term (according to p. 421, *Atombau*) downwards by the amount

$$\frac{1}{4} \frac{\alpha^2 R}{n^4},$$

and on the other hand the "separations" (*Aufspaltungen*) $\Delta\nu$, $\Delta\nu_1$, $\Delta\nu_2$. . . for $n = 2, 3, \dots$ and $k < n$. In the right-hand side of the figure we show how the position of these levels would be displaced by the new formula (47). *The electron spin* (or, as we may also say, *the electron magnetism*) must cancel this displacement and restore the levels to the positions occupied in the middle column of the diagram. We have used arrows to indicate how this restoration takes place: every level with $l > 0$ separates into two parts and moves partly upwards and partly

* We shall, however, see in Chap. II, § 10, that this objection can be overcome by giving the "boundary conditions" a broader basis.

downwards; only the levels with $l = 0$ move *only* upwards. It is on this account that all levels except the uppermost in every fine-structure configuration are to be counted as double and belong to two neighbouring values of l . This brings us to a notation for the levels, which corresponds

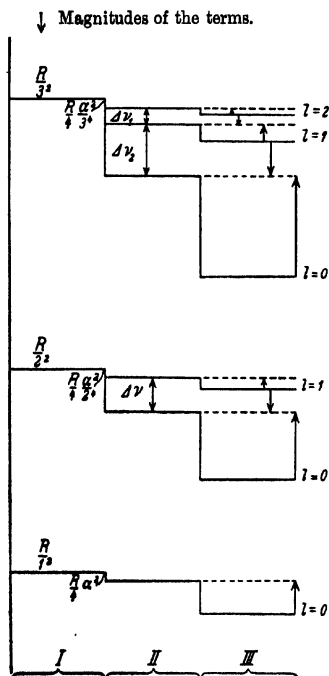


FIG. 10.

without re- old relativ- new relativ-
lativity. istic form. istic form.

Fine structure of the hydrogen terms. Only the elements of each term in Columns II and III have been drawn to scale; the distances in Column I and the relative measures of the different terms have been chosen arbitrarily. The new relativistic formula is wrong, as it does not take into account the "electron spin." The correct relativistic formula in Chap. II, § 10, will confirm the levels of Column II.

to the alkali character of the hydrogen spectrum emphasised in the previous section and which is made precise by the quantum number j there already introduced.

We illustrate the further details in Fig. 11 that corresponds to the hydrogen line H_α . Those levels, which are in reality coincident, are separated a little in the figure. On the left is the notation according to the old theory, the quantum number k being used; on the right is the new notation, which uses the numbers l and $j = l \pm \frac{1}{2}$, in which (cf. p. 93) $\pm \frac{1}{2}$ allows us to discern the action of the electron spin. Further, the term symbols s, p, d , which are customarily used in the spectra of the alkalis, have been added; in front of these values (on the extreme left) are the values of the principal quantum numbers n : the term symbols have also a suffix which should really be j , that is, half-integral but is here expressed integrally as $j + \frac{1}{2}$.

The difference between the old and the new notation finds expression in the number of allowed combinations. According to the old notation there are three (cf. *Atombau*, Chap. VI, p. 432, and I, Chap. VIII, p. 485, not only for H_α line but for all the Balmer lines).

According to the new notation we have seven possible transitions in all, three of the character (pd), two each of the character (ps) or (sp). These seven transitions actually occur in the alkali and Röntgen terms, and are all different; the notation of the resulting Röntgen lines (L-series) is indicated in the middle of the figure by means of the extended Moseley

notation which we used formerly; below it is the Siegbahn notation (cf. *Atombau*, Chap. IV, p. 284). In the case of hydrogen, however, two pairs of these seven transitions become identical as they have the same initial and final levels; they have been bracketed together in the figure at the bottom edge. There thus remain five fine-structure components for H_α and likewise for all the remaining Balmer lines. The two components that have been added, in comparison with the old theory, are indicated by arrows at the bottom of the figure. Although they are weak they have actually been found by Hansen*—without a knowledge of the new theory—by analysing his photometric photographs.

The same is true of the other hydrogen series. The *Lyman series*, which, according to the earlier theory was to consist of "strictly simple lines" (cf. I, Chap. VIII, p. 487) is, from the new point of view, a doublet series like the principal series of the alkalis. The Paschen series and the corresponding Fowler series of He^+ , with its famous line $\lambda = 4686$, were earlier to consist of five lines (cf. I, p. 490); the new view, based on the similarity with the series of the alkalis, leads in all to 13 components [namely, $3 + 3 + 3 + 2 + 2$ transitions of the character (df), (pd), (dp), (sp) and (ps)] of which five pairs, however, coincide, so that eight different components remain, that is, three more than according to the old theory: Paschen's photographs I, Figs. 114a-117, pp. 489-493, give, in part, direct evidence of these three components, as was first noticed by Goudsmit and Uhlenbeck, whereas earlier their appearance had to be ascribed to an incipient Stark effect, which was an unsatisfactory reason.

The fact that in the photographs of Försterling and Hansen the Balmer lines exhibit the Paschen-Back effect was mentioned on p. 94 as a convincing argument in favour of regarding the hydrogen spectrum as of the alkali type.

F. Historical and Critical Remarks

The relativistic generalisation of the wave-equation was given by Schrödinger in his fourth paper on "Quantisation as a Problem of Proper Values" and almost simultaneously found by various other authors, of whom we mention—

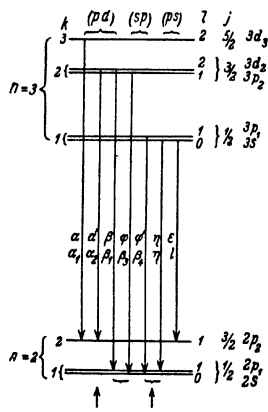


FIG. 11.—New quantum notation for H_α and comparison with alkali and Röntgen terms.

* Ann. d. Phys., 78, 558 (1925).

O. Klein, *Zeitschr. f. Phys.*, **37**, 895 (1926).

V. Fock, *ibid.*, **38**, 242 (1926); **39**, 226 (1926).

J. Kudar, *Ann. d. Phys.*, **81**, 632 (1926).

W. Gordon, *Zeitschr. f. Phys.*, **40**, 117 (1926).

Th. de Donder and H. van den Dungen, *Comptes Rendus*, July, 1926.

The relativistic generalisation of the variation principle is developed in the papers by Gordon and Fock just quoted. The relativistic four-current, besides being defined by Gordon *loc. cit.*, is also defined by Schrödinger (*Collected Papers*, p. 130).

Also, the new view of the hydrogen spectrum was proposed by several physicists—

S. Goudsmit and G. E. Uhlenbeck, *Physica*, **5**, 266 (1925).

J. C. Slater, *Proc. Nat. Acad.*, **11**, 732 (1925).

A. Sommerfeld and A. Unsöld, *Zeitschr. f. Phys.*, **36**, 259 (1926); **38**, 237 (1926).

Finally, we must call attention to the fact that the most recent papers by Dirac, in the *Proc. Roy. Soc.* of February and March, 1928, have given the relativistic generalisation of the wave-equation a totally new aspect. Dirac shows that this problem is very intimately connected with the gyroscopic nature of the electron, and that the electron spin or the electron magnetism is a necessary analytical consequence of the relativistic wave-equation, when properly formulated. We shall go into this in the next chapter, and shall show in what points the theory developed in this section must be altered in order to correspond with the new discoveries of Dirac.

This discovery settles in an extremely satisfactory way the time-worn controversy about the meaning of the relativistic fine structure. Originally derived from the relativity theory the fine-structure formula was claimed by various writers to be of magnetic origin. Millikan and Bowen* believed themselves compelled to conclude from the great amount of information gained by them about "stripped atoms" that the "splitting up" of the levels must take place relativistically only in the case of hydrogen (He^+ and so forth), but in the case of the other atoms and the Röntgen spectra this separation was conditioned magnetically; this made the identity of the formula which was valid in both cases appear to be due to chance. Through Dirac we now know that this identity is rooted in the nature of the electron magnetism. This electron magnetism is itself relativistic in origin. Instead of an incomprehensible accident we see here the manifestation of a deep-seated identity which must ultimately lead to a solution of the problem of the electron itself, which includes such questions as: why do negative and positive electricity occur in elementary quanta, why is their mass unequal, what is their structure, how do they hold together?

* *Phil. Mag.*, **49**, 928 (1925).

§ 10. The Zeeman Effect

The normal Zeeman effect, as first theoretically explained by H. A. Lorentz occurs, according to our modern view, only in the case of true singlet lines. These are due to transitions between two states in each of which the resultant electron spin s (cf. p. 86) compensates itself and becomes zero; for example, parhelium with its two electrons spinning in opposite directions. The hydrogen lines for which we formerly conjectured that the normal Zeeman effect would occur do not belong to the singlets and in reality show the anomalous Zeeman effect of the alkalis (cf. § 9, E). Our present basis of wave-mechanics is sufficient to treat the *normal Zeeman effect* fully and without difficulty. But for the *anomalous Zeeman effect* it needs to be supplemented by the spin of the electron.

The wave-mechanical treatment of the *normal Zeeman effect* is founded on eqn. (10) of § 9. This equation neglects the relativistic corrections, which are of no interest for the Zeeman effect, but it gives expression to the influence of the magnetic field by means of the term $\mathbf{A} \text{ grad } u$. In the case of a homogeneous magnetic field \mathbf{H} parallel to the z axis we set, as in *Atombau*, Note 7, p. 802,

$$\mathbf{A}_x = -\frac{1}{2}Hy, \quad \mathbf{A}_y = \frac{1}{2}Hx, \quad \mathbf{A}_z = 0 \quad . \quad . \quad . \quad (1)$$

By the formula $\mathbf{H} = \text{curl } \mathbf{A}$ we then get

$$\mathbf{H}_x = \mathbf{H}_y = 0, \quad \mathbf{H}_z = H.$$

From (1) we calculate

$$(\mathbf{A} \text{ grad } u) = \frac{H}{2} \left(x \frac{\partial u}{\partial y} - y \frac{\partial u}{\partial x} \right) \quad . \quad . \quad . \quad (2)$$

But the expression in brackets on the right is equal to $\frac{\partial u}{\partial \phi}$ if we introduce polar co-ordinates r, θ, ϕ , as may easily be verified. So we also have

$$(\mathbf{A} \text{ grad } u) = \frac{H}{2} \frac{\partial u}{\partial \phi},$$

and eqn. (10) in § 9 transforms into

$$\Delta u + \frac{4\pi i}{h} \left\{ \frac{eH}{2c} \frac{\partial u}{\partial \phi} - \mu \frac{\partial u}{\partial t} \right\} - \frac{8\pi^2 \mu}{h^2} (E_0 + U)u = 0 \quad . \quad . \quad (3)$$

To prevent confusion with the magnetic quantum number m we have written μ for m_0 . To integrate this we set

$$u = \psi e^{\frac{2\pi i E_0 t}{h}}, \quad \psi = R P_l^m(\cos \theta) e^{im\phi} \quad . \quad . \quad . \quad (4)$$

The exponential form for the dependence on ϕ is necessary here, and is not merely dictated by considerations of convenience as in the Kepler problem for the case where no magnetic field is acting. For, on account of the term with $\frac{\partial u}{\partial \phi}$ we should not be able to solve eqn. (3) at all by

separation of the variables, if we were to write $\frac{\cos}{\sin} m\phi$ instead of $e^{im\phi}$. This remark is essential for the arguments in § 8, p. 92, which justified our calculation of the magnetic moment M . We here encounter for the first time a case for which the complex form of ψ is conditioned by the nature of the problem.

If we substitute (4) in (3), we get as the differential equation for the dependence of ψ on r and θ

$$\Delta\psi + \frac{8\pi^2\mu}{h^2} \left\{ E' - \frac{heH}{4\pi\mu c} m - E_0 - U \right\} \psi = 0 \quad (5)$$

This equation becomes identical with eqn. (1) of § 7 if we make

$$E' - \frac{heH}{4\pi\mu c} m - E_0 = E \quad (6)$$

Consequently all the results of § 7 concerning the proper values E and the proper functions ψ apply here: *the presence of a magnetic field causes the states (Zustände) to differ only in the value of the energy, not in the form of the proper functions.* This is the wave-mechanical analogy to Larmor's theorem. If we also introduce the Larmor precession (I, Chap. V, § 6, eqn. (2), p. 296)

$$\omega_L = \frac{1}{2} \frac{e}{\mu} \frac{H}{c} \quad (6a)$$

eqn. (6) becomes

$$E' = E_0 + E + \frac{h}{2\pi} \omega_L m \quad (7)$$

If, further, we introduce the circular frequency ω of the vibration represented by (4), by setting

$$e^{\frac{2\pi i}{h} E' t} = e^{i\omega t},$$

we get, by (7),

$$\omega = \frac{2\pi}{h} (E_0 + E) + \omega_L m \quad (8)$$

and the exponential functions in (4) combine to

$$e^{i(\omega t + m\phi)} \quad (9)$$

Accordingly, the state may be described as a *circular vibration*, which takes place around the direction of the magnetic lines of force. The phase of the vibration is $\omega t + m\phi$, and hence the phase velocity, defined as in eqn. (4), § 5, is

$$a = \frac{d\phi}{dt} = - \frac{\omega}{m} \quad (10)$$

Since $|E| \ll E_0$ and $\omega_L m$ is therefore so much the more $\ll E_0$, we get from (8)

$$\omega \sim \frac{2\pi}{h} E_0 = \frac{2\pi\mu}{h} c^2 \sim 9 \cdot 10^{20} \text{ sec}^{-1}$$

and hence from (10)

$$a \sim -\frac{1}{m} 9 \cdot 10^{20} \text{ sec}^{-1}.$$

Thus the state runs round in the circle at an enormous rate; the period of vibration is considerably smaller than that of hard Röntgen rays. This motion in the circle appears to stand in no relation to the classical Larmor precession.

This has led Epstein* to conclude that the very simple and direct treatment of the Zeeman effect just given is incorrect and must be supplanted by a more detailed discussion, in which the magnetically influenced atom must be combined with the current producing the magnetic field into one system, as was done by H. A. Lorentz. There is no doubt that Epstein's method is instructive and well founded physically, but the following consideration† shows that our more direct method is justified.

If we compare our expression (9) for the progressive circular wave with the expression (2) in § 5 for the progressive linear wave, we recognise that m takes the place of the earlier wave-number k . For, actually, m denotes the number of waves that fall within the angle 2π , just as k denotes the number of waves that occur in the length 2π . Thus if we wish to calculate the group velocity b that is associated with the phase velocity a , by (7) of § 5, we must form

$$b = \frac{d\omega}{dm}.$$

The differential quotient on the right is to be taken from eqn. (8), which, so to speak, gives the law of dispersion for our circular wave, that is, the dependence of frequency on the "wave-number" m . This yields, since E_0 and E are independent of m ,

$$b = \frac{d\omega}{dm} = \omega_L \quad . \quad . \quad . \quad . \quad (11)$$

Thus whereas the phase velocity is immensely great, the group velocity comes out equal to the well-known slow angular velocity of the Larmor precession.

This evident and striking relationship with the Larmor precession not only accounts for the surprisingly high value of the phase velocity but serves as a subsequent justification of the preceding method. The conditions here are quite analogous to those which occur in the de Broglie waves: for then, by § 5, the wave-phase propagates itself with a velocity greater than that of light (*Überlichtgeschwindigkeit*, $\frac{c^2}{v}$), but the wave-group moves with only the velocity v of matter. It suggests itself immediately to draw inferences in the present case, too, from the group velocity to the motion of matter, that is, to assume that the Larmor precession is now

* Nat. Acad. Washington, 12, 684, November, 1926.

† I am indebted for this to a lecture given in 1927 by E. U. Condon.

also to be interpreted as a circling motion of the electrons, as in classical mechanics.

But we must point out an apparent contradiction. In eqn. (11) we differentiated with respect to the quantum number m , as if we were dealing with a continuously variable quantity, whereas by the postulate that ψ shall be single-valued, m is fixed as an integer. To resolve this contradiction we must refer to the general remarks at the beginning of § 8. We there stated: if the energy is sharply defined as a proper value, the time co-ordinate of the electron becomes indefinite and the idea of an orbit becomes diffuse; so we must give up the sharp definition of the energy. By eqn. (7) m occurs in the energy E' ; by varying m continuously, E' is also continuously varied. This is necessary if we are to be able to speak of a group velocity at all as the velocity of an electron in its orbit. We thus see that the apparently inadmissible differentiation with respect to m is connected with the Uncertainty Relation mentioned in § 8, A (*Ungenauigkeitsrelation*).

Moreover, this same difficulty occurs even in the rectilinear motion of the electron. Here, too, according to de Broglie, the velocity v of the electron is obtained by differentiating with respect to the wave-length λ (or to its reciprocal, the wave-number k). λ is capable of having continuous values so long as we consider the electron in unlimited space. But as soon as we imagine it enclosed in a cavity (Hohlraum), the λ 's are in this case also fixed as discrete values, analogously to our m in the Zeeman effect. The allocation of the group velocity v to the electron wave here also then requires the rescinding of this restriction to fixed values and renders it necessary to ascribe to the wave-length λ , just as to our m , sufficient play for a continuous series of values.

We can now easily show that our wave-mechanical treatment of the normal Zeeman effect contains all the results that we derived in I, Chap. V, § 6, from the earlier theory. By eqn. (7) the additional energy ΔE due to the magnetic field, that is, the difference between E' and $E_0 + E$ is given by

$$\Delta E = \frac{m}{2\pi} \omega_L h.$$

This is (if we disregard slight differences in notation) identical with I, eqn. (12), p. 299. By taking the difference between an initial state 1 and a final state 2, we obtain the magnetic change of wave-number ν by the frequency condition

$$\Delta \nu = \frac{\Delta E_1 - \Delta E_2}{h} = \frac{m_1 - m_2}{2\pi} \omega_L,$$

or, taking into consideration (6a)

$$\Delta \nu = (m_1 - m_2) \frac{e}{\mu} \frac{H}{4\pi c} \quad (12)$$

This is the same as the earlier equation (14) on p. 300 of Vol. I.

The superiority of the new method over the old manifests itself, however, in the fact that we now obtain also the rules of selection and polarisation from the same mathematical scheme as the energy values. Since the proper functions and the "matrix elements" of the co-ordinates, which are formed from them, are the same as those in the Kepler problem when no magnetic field is acting, we can invoke eqn. (37a) of § 7, E, for the rule of selection for m . This states that only the transitions

$$m \rightarrow m \pm 1$$

may occur with an intensity differing from zero. At the same time we can take the corresponding polarisations from § 6, B, p. 55. According to this, for a transition $m \rightarrow m$ a vibration arises which is polarised linearly along the z -axis (= that of the magnetic field), and for a transition $m \rightarrow m \pm 1$ a circular vibration occurs in each case in the plane perpendicular to the magnetic field. In the first case (linear polarisation) $m_1 - m_2 = 0$ and by eqn. (12) $\Delta\nu = 0$; in the second case (circular polarisation) $m_1 - m_2 = \pm 1$, and so, by (12),

$$\Delta\nu = \pm \frac{e}{\mu} \frac{H}{4\pi c}.$$

In this way we have derived the *normal Lorentz triplet*, corresponding to I, eqn. (16), p. 300, and I, Fig. 81, p. 294. It is evident that if we now evaluate the matrix elements in question we also obtain the intensities of the three triplet-components, namely, in the case of transverse observation, twice as great an intensity for the middle component as for either of the lateral components.

Finally, we must point out again that the *normal Zeeman effect* is to be expected only under quite special conditions, and that even in the case of the H-atom the Zeeman effect that occurs is anomalous.

§ 11. The Molecule as a Symmetrical Top

In I, Chap. VII, p. 441, we distinguish between *diatomic molecules*, whose moment of inertia about the line connecting the nuclei vanishes and which can therefore take up no moment of momentum about the axis, and molecular tops (Kreisel-Molekülen), whose ellipsoid of inertia is an ellipsoid of rotation and which can therefore have a spin-moment about the axis of symmetry. The latter correspond to the "symmetrical tops" of ordinary mechanics, so long as we regard the configuration of atoms as rigid. The much more complicated problem of the unsymmetrical top (general ellipsoid of inertia) was only hinted at earlier (*Atombau*, p. 744) and will not be treated in this volume either.

The question of the quantising of the symmetrical molecular top has gained in interest since Victor Henri* discovered in formaldehyde

* *Nature*, 118, 225 (1926). Victor Henri and Svend Aage Schou, *Zeitschr. f. Phys.*, 48, 774, 1928.

(CH₂O) the case of a substance whose band spectra exhibit the unmistakable characteristics of *both* moments of inertia, namely C about the axis of the figure of the molecule and A about the equatorial axis. The quantum-formula for such molecules was derived earlier in I, Chap. 7, § 6, eqn. (7), p. 443. The wave-mechanical method of derivation confirms this formula but again with the characteristic difference that $m(m+1)$ takes the place of the m^2 of the earlier theory (where m = the rotation quantum number for the total angular momentum). We do not here consider the oscillation terms of the band spectrum as we decided to treat the molecules as rigid bodies.

We enter into this problem rather deeply, not only on account of its importance for the theory of band spectra, but also on account of the method involved. On the one hand, it gives us an opportunity of introducing generalised co-ordinates, here the Eulerian angles, into the wave-equation; on the other hand, it gives us an excellent example of the power of our "polynomial method," by means of which we shall simplify the calculation as compared with the other methods that have appeared.*

A. The Wave Equation in Generalised Co-ordinates

Just as in ordinary mechanics Hamilton's principle, so in wave-mechanics the variation principle of § 9, D, simplifies the introduction of generalised co-ordinates. As we are not here interested in relativity corrections we may use the simplified form (38a) of the variation principle.

*As in (38a) we consider first the single point-mass and express its three rectilinear co-ordinates x, y, z in terms of three generalised co-ordinates θ, χ, ϕ , or in general by q without specifying them any further. We then have

$$\frac{\partial \psi}{\partial x} = \frac{\partial \psi}{\partial \theta} \theta_x + \frac{\partial \psi}{\partial \chi} \chi_x + \frac{\partial \psi}{\partial \phi} \phi_x,$$

$$\frac{\partial \psi}{\partial y} = \frac{\partial \psi}{\partial \theta} \theta_y + \dots,$$

in which the suffixes denote differentiations. Using the abbreviations

$$\left. \begin{aligned} [\theta\theta] &= \frac{1}{2m}(\theta_x^2 + \theta_y^2 + \theta_z^2) \\ [\theta\chi] &= \frac{1}{2m}(\theta_x\chi_x + \theta_y\chi_y + \theta_z\chi_z) \dots \end{aligned} \right\} \quad (1)$$

* F. Reiche, *Zeitschr. f. Phys.*, **39**, 444 (1926); R. de L. Kronig and Rabi, *Phys. Rev.*, **29**, 262 (1927); O. Mannebeck, *Phys. Zeitschr.*, **28**, 72 (1927); D. M. Dennison, *Phys. Rev.*, **23**, 818 (1926), in which the method of matrices is used. Concerning questions of intensity, cf. Rademacher and Reiche, *Zeitschr. f. Phys.*, **41**, 458 (1927).

we then have for the differential expression that occurs in eqn. (38a) of § 9

$$\left. \begin{aligned} \frac{1}{2m} \left[\left(\frac{\partial \psi}{\partial x} \right)^2 + \left(\frac{\partial \psi}{\partial y} \right)^2 + \left(\frac{\partial \psi}{\partial z} \right)^2 \right] \\ = [\theta\theta] \left(\frac{\partial \psi}{\partial \theta} \right)^2 + [\chi\chi] \left(\frac{\partial \psi}{\partial \chi} \right)^2 + \dots + 2[\theta\chi] \frac{\partial \psi}{\partial \theta} \cdot \frac{\partial \psi}{\partial \chi} + \dots \end{aligned} \right\} \quad (2)$$

But the same coefficients $[\theta\theta]$ $[\theta\chi]$. . . also occur in the expression for the *vis viva* if we write it as a function of the generalised co-ordinates of position q and momenta $p = p_\theta, p_\chi, p_\phi$. For we have

$$\left. \begin{aligned} \dot{\theta} &= \theta_x \dot{x} + \theta_y \dot{y} + \theta_z \dot{z} \\ \dot{\chi} &= \chi_x \dot{x} + \dots \end{aligned} \right\} \quad (3)$$

and hence

$$\left. \begin{aligned} \frac{\partial T}{\partial \dot{x}} &= \frac{\partial T}{\partial \dot{\theta}} \frac{\partial \dot{\theta}}{\partial \dot{x}} + \frac{\partial T}{\partial \dot{\chi}} \frac{\partial \dot{\chi}}{\partial \dot{x}} + \frac{\partial T}{\partial \dot{\phi}} \frac{\partial \dot{\phi}}{\partial \dot{x}} = \theta_x \frac{\partial T}{\partial \dot{\theta}} + \chi_x \frac{\partial T}{\partial \dot{\chi}} + \phi_x \frac{\partial T}{\partial \dot{\phi}} \\ \frac{\partial T}{\partial \dot{y}} &= \theta_y \frac{\partial T}{\partial \dot{\theta}} + \dots \end{aligned} \right\} \quad (4)$$

According to the general definition of momentum co-ordinates (I, Chap. IV, § 1, eqn. 5, p. 195) the left-hand sides of these equations are the rectilinear momenta p_x, p_y, p_z , and the differential quotients on the right are the momenta $p_\theta, p_\chi, p_\phi^*$. If we substitute these terms in (4) we get

$$\begin{aligned} \dot{x} &= \frac{1}{m} (\theta_x p_\theta + \chi_x p_\chi + \phi_x p_\phi) \\ \dot{y} &= \frac{1}{m} (\theta_y p_\theta + \dots) \end{aligned}$$

Hence, taking into account (1), we get

$$\left. \begin{aligned} T(q, p) &= \frac{m}{2} (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) \\ &= [\theta\theta] p_\theta^2 + [\chi\chi] p_\chi^2 + \dots + 2[\theta\chi] p_\theta p_\chi + \dots \end{aligned} \right\} \quad (5)$$

We actually have here, then, the same coefficients $[\theta\theta]$. . . as in (2) and so we read out of (5) the following rule, already formulated on p. 111, but now extended to generalised co-ordinates: *the expression H in eqn. (38a), § 9, is formed by multiplying in turn each term of the Hamiltonian function of classical mechanics by ψ^2 and replacing $p_k \psi$ by*

$$\frac{h}{2\pi} \frac{\partial \psi}{\partial q_k}$$

* We may remark parenthetically that a comparison of (3) and (4) exhibits the well-known fact that momenta transform contravariantly with respect to the velocities.

This gives the transformed expression for H

$$H = T\left(q, \frac{h}{2\pi} \frac{\partial \psi}{\partial q}\right) + V\psi^2 = \frac{h^2}{4\pi^2} \left\{ [\theta\theta] \left(\frac{\partial \psi}{\partial \theta}\right)^2 + [\chi\chi] \left(\frac{\partial \psi}{\partial \chi}\right)^2 + \dots + 2[\theta\chi] \frac{\partial \psi}{\partial \theta} \frac{\partial \psi}{\partial \chi} + \dots \right\} + V\psi^2 \quad (6)$$

The variation principle of (38a, b), § 9 then runs, if as on p. 111 we call the Lagrange multiplier of the normalising condition $-E$,

$$\delta \int \left\{ T\left(q, \frac{h}{2\pi} \frac{\partial}{\partial q}\right) - E + V \right\} \psi^2 d\tau = 0 \quad (7)$$

Of course $d\tau = dx dy dz$ must also be transformed correspondingly, that is, must be replaced by $D d\theta d\chi d\psi$, abbreviated to $D dq$, where D is the functional determinant of the transformation :

$$D = \begin{vmatrix} x_\theta & y_\theta & z_\theta \\ x_\chi & y_\chi & z_\chi \\ x_\psi & y_\psi & z_\psi \end{vmatrix} \quad (8)$$

The indices here denote differentiations, as above :

$$x_\theta = \frac{\partial x}{\partial \theta} \text{ and so forth}$$

corresponding to the earlier

$$\theta_x = \frac{\partial \theta}{\partial x} \dots$$

Performing the variation in (7), we next get

$$\int \left\{ D \sum_k \frac{\partial T}{\partial p_k} \cdot \frac{h}{2\pi} \frac{\partial \delta \psi}{\partial q_k} - 2D(E - V)\psi \delta \psi \right\} dq = 0$$

and after an integration by parts

$$- \int \left\{ \frac{h}{2\pi} \sum_k \frac{\partial}{\partial q_k} \left(D \frac{\partial T}{\partial p_k} \right) + 2D(E - V)\psi \right\} \delta \psi dq = 0 \quad (9)$$

The "surface integrals" resulting from the integration by parts have been omitted, as they can be made to vanish, if necessary, by the subsidiary condition $\delta \psi = 0$ "at the boundary." Since $\delta \psi$ is moreover arbitrary, the factor of $\delta \psi$ in (9) must be equated to zero. This gives the *transformed wave-equation*. In view of eqn. (5), in which p_k must now of course be replaced by

$$\frac{h}{2\pi} \frac{\partial \psi}{\partial q_k},$$

it runs, explicitly

$$\left. \begin{aligned} & \frac{\partial}{\partial \theta} \left\{ D \left([\theta \theta] \frac{\partial \psi}{\partial \theta} + [\chi \theta] \frac{\partial \psi}{\partial \chi} + [\phi \theta] \frac{\partial \psi}{\partial \phi} \right) \right\} \\ & + \frac{\partial}{\partial \chi} \left\{ D \left([\theta \chi] \frac{\partial \psi}{\partial \theta} + [\chi \chi] \frac{\partial \psi}{\partial \chi} + [\phi \chi] \frac{\partial \psi}{\partial \phi} \right) \right\} \\ & + \frac{\partial}{\partial \phi} \left\{ D \left([\theta \phi] \frac{\partial \psi}{\partial \theta} + [\chi \phi] \frac{\partial \psi}{\partial \chi} + [\phi \phi] \frac{\partial \psi}{\partial \phi} \right) \right\} \\ & + \frac{4\pi^2}{h^2} D(E - V) \psi = 0 \end{aligned} \right\} \quad (10)$$

It is self-adjoint [cf. the conditions* (19) in § 9] like every differential equation that arises from a variation principle.

B. Application to Euler's Angles and to the Rotation of a Rigid Body

The form of equation just obtained holds not only for the individual point-mass but equally well for a system of point-masses; we have but to increase the number of co-ordinates, that is, to make them equal to the number of degrees of freedom of the system. We have already seen in eqn. (12), § 1, that in the case of several point-masses (distinguished by the index α), it was only necessary to replace $\Delta\psi/m$ in the wave-equation by

$$\sum_{\alpha} \frac{1}{m_{\alpha}} \Delta_{\alpha} \psi.$$

From this it follows that in the variation principle, eqn. (38a), § 9, the expression

$$\frac{1}{2m} \left[\left(\frac{\partial \psi}{\partial x} \right)^2 + \dots \right]$$

must be replaced by a corresponding sum taken over α with m_{α} in the denominator. But since, by (5), T can also be represented by an analogous sum, (6) and (10) remain unchanged in form, $[\theta\theta]$, $[\theta\chi]$, . . . now of course signifying the coefficients of $T(q, p)$ for the whole point system.

The point system of the rigid body has three degrees of freedom; Euler's angles are three appropriate co-ordinates, and are usually called θ , ψ , ϕ , but as we have claimed ψ for the wave-function we shall retain the notation θ , χ , ϕ used in A of this section. These angles, as we know, define the position of the fixed system of reference of the axis of the figure in the case of the top (the Z -axis = the axis of rotation of the ellipsoid of inertia) and of the equatorial plane (XY -plane) with reference to an xyz -system fixed in space but otherwise arbitrary in position; the line of intersection of the XY -plane with the xy -plane is called the line of *nodes* (*Knotenlinie*), θ denotes the angle between the z and the Z -axis, χ the

* These conditions are, it is true, formulated for only two independent variables x, y . But we can apply them to each pair of the variables θ, χ, ϕ in turn and show that they are fulfilled.

angle between the line of nodes and the X axis; θ lies between 0 and π , χ between 0 and 2π .

The kinetic energy of the symmetrical top is given by a function of these angles and of the corresponding angular velocities

$$T(q, \dot{q}) = \frac{A}{2}(\dot{\theta}^2 + \sin^2 \theta \dot{\chi}^2) + \frac{C}{2}(\dot{\phi} + \cos \theta \dot{\chi})^2 \quad (11)$$

To avoid leaving a gap here we prove this formula very simply as follows. We start from the general equation of definition of the principal moments of inertia

$$T = \frac{1}{2}(A\omega_x^2 + B\omega_y^2 + C\omega_z^2);$$

$\omega_x, \omega_y, \omega_z$, refer to the above-mentioned principal axes X, Y, Z, fixed in the top (we avoid the usual notation p, q, r , because we have already used

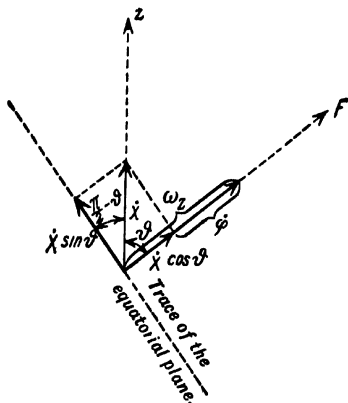


FIG. 12a.—Plane through the axis F of the figure (Z-axis) and the x -axis. Projection of the rotation vector on the axis F and on the trace of the equatorial plane.

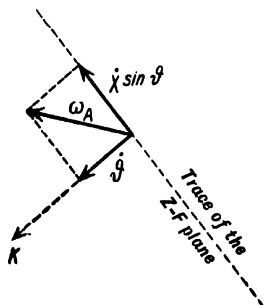


FIG. 12b.—Projection of the rotation vector on the equatorial plane.

p and q for other purposes). Hence for the symmetrical top $A = B$ we have

$$T = \frac{A}{2}\omega_{\perp}^2 + \frac{C}{2}\omega_z^2 \quad (11a)$$

$\omega_{\perp} = (\omega_x^2 + \omega_y^2)^{1/2}$ denotes the perpendicular projection of the rotation vector ω on the equatorial plane, ω_z that on the axis of the figure. Besides these components we consider the *oblique* components of the rotation vector, namely $\dot{\theta}$, $\dot{\chi}$, $\dot{\phi}$ along the line of nodes, the z -axis, and the axis of the figure. Thus we have resolved the rotation vector ω in two ways, once in the rectangular components ω_z and ω_{\perp} , and secondly in the oblique complex of lines $\dot{\theta}$, $\dot{\chi}$, $\dot{\phi}$ (cf. Fig. 12a, b). If we project this set of lines on the axis of the figure, we get ω_z . Thus (Fig. 12a)

$$\omega_z = \dot{\phi} + \cos \theta \dot{\chi} \quad (11b)$$

θ makes no contribution to ω_z , since the line of nodes is perpendicular to the axis of the figure. We then project this set of lines on the equatorial plane (Fig. 12b). In this case ϕ makes no contribution because the axis of the figure is perpendicular to the equatorial plane. θ lies of itself in the equatorial plane, namely along the line of nodes K. The projection of χ lies in the direction perpendicular to the line of nodes (namely in the trace of the plane of the diagram in Fig. 12a). So we get

$$\omega_A^2 = \theta^2 + \sin^2 \theta \chi^2.$$

If we now substitute (11b, c) in (11a) we find that we have proved (11).

From (11) it follows that

$$p_\theta = A\dot{\theta}, \quad p_\chi = (A \sin^2 \theta + C \cos^2 \theta)\dot{\chi} + C \cos \theta \dot{\phi}, \\ p_\phi = C(\dot{\phi} + \cos \theta \dot{\chi}).$$

By solving and substituting in T (q, \dot{q}) we get

$$T(q, p) = \frac{p_\theta^2}{2A} + \frac{(p_\chi - \cos \theta p_\phi)^2}{2A \sin^2 \theta} + \frac{p_\phi^2}{2C} \quad (11d)$$

Accordingly,

$$\left. \begin{aligned} [\theta\theta] &= \frac{1}{2A}, & [\chi\chi] &= \frac{1}{2A \sin^2 \theta}, & [\phi\phi] &= \frac{\cos^2 \theta}{2A \sin^2 \theta} + \frac{1}{2C} \\ [\theta\chi] &= [\theta\phi] = 0, & [\chi\phi] &= -\frac{\cos \theta}{2A \sin^2 \theta} \end{aligned} \right\} \quad (12)$$

It still remains to determine the value of the functional determinant D, eqn. (8). To refer it to the preceding quantities, we may proceed as follows. Besides D we consider the determinant

$$D_1 = \begin{vmatrix} \theta_x & \theta_y & \theta_z \\ \chi_x & \chi_y & \chi_z \\ \phi_x & \phi_y & \phi_z \end{vmatrix} \quad (13)$$

as well as its square and its product with D. The last two quantities are to be formed according to the multiplication theorem for determinants. D_1^2 is then equal to the "discriminant" of the quadratic form (11d), namely,

$$D_1^2 = \begin{vmatrix} [\theta\theta] & [\theta\chi] & [\theta\phi] \\ [\chi\theta] & [\chi\chi] & [\chi\phi] \\ [\phi\theta] & [\phi\chi] & [\phi\phi] \end{vmatrix} = \begin{vmatrix} \frac{1}{2A} & 0 & 0 \\ 0 & \frac{1}{2A \sin^2 \theta} & -\frac{\cos \theta}{2A \sin^2 \theta} \\ 0 & -\frac{\cos \theta}{2A \sin^2 \theta} & \frac{\cos^2 \theta}{2A \sin^2 \theta} + \frac{1}{2C} \end{vmatrix} \quad (14) \\ = \frac{1}{8A^2 C \sin^2 \theta}$$

On the other hand, we easily prove that

$$DD_1 = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1 \quad (15)$$

on account of the relationships

$$\begin{aligned}\theta_x x_\theta + \theta_y y_\theta + \theta_z z_\theta &= \left(\frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial y} dy + \frac{\partial \theta}{\partial z} dz \right) / d\theta \\ &= \left(\frac{d\theta}{d\theta} \right)_{x\theta} = 1 \text{ and so forth.} \\ \theta_x x_\chi + \theta_y y_\chi + \theta_z z_\chi &= \left(\frac{\partial \theta}{\partial x} dx + \frac{\partial \theta}{\partial y} dy + \frac{\partial \theta}{\partial z} dz \right) / d\chi \\ &= \left(\frac{\partial \theta}{\partial \chi} \right)_{\theta\chi} = 0 \text{ and so forth.}\end{aligned}$$

Hence from (14) and (15) we get

$$D = \frac{1}{D_1} = 2A \sin \theta \sqrt{2C} \quad . \quad . \quad . \quad (16)$$

We now need only substitute the values (12) and (16) in the scheme (10) to obtain the wave equation of our problem

$$\left. \begin{aligned} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin \theta} \frac{\partial^2 \psi}{\partial \chi^2} + \left(\frac{\cos^2 \theta}{\sin \theta} + \frac{A}{C} \sin \theta \right) \frac{\partial^2 \psi}{\partial \phi^2} \\ - \frac{2 \cos \theta}{\sin \theta} \frac{\partial^2 \psi}{\partial \chi \partial \phi} + \frac{8\pi^2 A}{h^2} \sin \theta (E - V) \psi = 0. \end{aligned} \right\} \quad (17)$$

As we shall be interested only in the motion of our molecule under no forces (not in its Stark effect, which has also been treated by Reiche) we must set $V = 0$, which means that E then signifies the difference between the actual energy and the energy at rest.

C. Integration of the Wave Equation

From (17) it next follows that χ and ϕ are cyclic co-ordinates, just as in the ordinary mechanics of the symmetrical top. In the latter this, as we know, causes the corresponding momenta p_χ and p_ϕ to be constant. In wave-mechanics the same circumstance allows us to assume the wave-function to depend on χ and ϕ in the special exponential form

$$\psi = \Theta(\theta) e^{i\tau\phi} + i\tau'\chi \quad . \quad . \quad . \quad (18)$$

Here τ and τ' must be integers (positive or negative), since ψ must be a *one-valued* function within the range of the co-ordinates and hence periodic in ϕ and χ with the period 2π . Eqn. (17) then becomes an ordinary differential equation in the unknown Θ

$$\left. \begin{aligned} \sin \theta \frac{d}{d\theta} \left(\sin \theta \frac{d\Theta}{d\theta} \right) - \left[\tau'^2 + \left(\cos^2 \theta + \frac{A}{C} \sin^2 \theta \right) \tau^2 \right. \\ \left. - 2 \cos \theta \tau \tau' - \frac{8\pi^2 A E}{h^2} \sin^2 \theta \right] \Theta = 0. \end{aligned} \right\} \quad (19)$$

As in the case of the spherical harmonics (*Kugelfunktionen*) we introduce the new independent variable $x = \cos \theta$ and use the abbreviation

$$\lambda = \frac{8\pi^2 A E}{h^2} + \tau^2 - \frac{A}{C} \tau^2 \quad . \quad . \quad . \quad (20)$$

From (19), by easy calculation, we then have

$$\left. \begin{aligned} (1-x^2)^2 \frac{d^2 \Theta}{dx^2} - 2x(1-x^2) \frac{d\Theta}{dx} \\ + [\lambda(1-x^2) - \tau^2 - \tau'^2 + 2\tau\tau'x]\Theta = 0 \end{aligned} \right\} \quad (21)$$

The only singular points of this differential equation in the range $-1 \leq x \leq +1$ are the limiting points $x = \pm 1$. To test the character of these singularities, we write

$$1 \mp x = y \quad (21a)$$

Thus

$$x = \pm(1-y), \quad 1 \pm x = 2-y \quad (21b)$$

and we have instead of (21), if we denote differentiations with respect to y by dashes,

$$\left. \begin{aligned} y^2(2-y)^2 \Theta'' + 2y(1-y)(2-y) \Theta' \\ + \{\lambda y(2-y) - (\tau \mp \tau')^2 \mp 2\tau\tau'y\} \Theta = 0 \end{aligned} \right\} \quad (22)$$

As in all the earlier cases we assume

$$\Theta = y^\gamma \Sigma \alpha_\nu y^\nu \quad (23)$$

and get the following characteristic equation for γ [by equating to zero the factor of y^γ in the power series of the left-hand side of (22)]:

$$4\gamma(\gamma-1) + 4\gamma - (\tau \mp \tau')^2 = 0$$

that is,

$$4\gamma^2 = (\tau \mp \tau')^2, \quad \gamma = \frac{|\tau \mp \tau'|}{2} \quad (24)$$

By introducing the vertical lines denoting absolute values we have fixed our choice of the roots, which is necessary to ensure the continuity of the proper function.

We further split off from Θ the two characteristic powers at the singular points; we count y , for example, from the point $x = -1$, so that by (21a, b) we get $1+x=y$, $1-x=2-y$. The assumption (23) then becomes changed to

$$\Theta = F \cdot v, \quad F = (2-y)^{\frac{|\tau-\tau'|}{2}} y^{\frac{|\tau+\tau'|}{2}} \quad (25)$$

We seek the differential equation that v satisfies, and for this purpose form

$$\begin{aligned} \Theta' &= F \left(v' - \frac{|\tau-\tau'|}{2(2-y)} v + \frac{|\tau+\tau'|}{2y} v \right), \\ \Theta'' &= F \left(v'' - \frac{|\tau-\tau'|}{2-y} v' + \frac{|\tau+\tau'|}{y} v' \right. \\ &\quad + \left\{ \frac{|\tau-\tau'|(|\tau-\tau'|-2)}{4(2-y)^2} + \frac{|\tau+\tau'|(|\tau+\tau'|-2)}{4y^2} \right. \\ &\quad \left. \left. - \frac{|\tau-\tau'|}{2y(2-y)} \frac{|\tau+\tau'|}{y} \right\} v \right). \end{aligned}$$

We substitute these values in eqn. (22), in which the lower sign must now be taken, and cancel the factor $y(2-y)F$. We get

$$y(2-y)v'' + \{ |\tau + \tau'| (2-y) - |\tau - \tau'| y + 2(1-y)v' \} + (\lambda + \Lambda)v = 0. \quad (26)$$

Λ is an abbreviation for the following expression :

$$\begin{aligned} \Lambda = & -\frac{(\tau + \tau')^2 + 2\tau\tau'y}{y(2-y)} + \frac{1}{2}(\tau - \tau')^2 \frac{y}{2-y} + \frac{1}{2}(\tau + \tau')^2 \frac{2-y}{y} \\ & - \frac{1}{2}|\tau^2 - \tau'^2| - \frac{1}{2}|\tau - \tau'| \left(\frac{y}{2-y} + \frac{2(1-y)}{2-y} \right) \\ & - \frac{1}{2}|\tau + \tau'| \left(\frac{2-y}{y} - \frac{2(1-y)}{y} \right). \end{aligned}$$

We show, however, that Λ is in reality a constant. For the first line reduces to $-\frac{1}{2}(\tau^2 + \tau'^2)$, and the two expressions in brackets in the second and third lines are equal to 1. Therefore

$$\Lambda = -\frac{1}{2}(\tau^2 + \tau'^2) - \frac{1}{2}|\tau^2 - \tau'^2| - \frac{1}{2}|\tau - \tau'| - \frac{1}{2}|\tau + \tau'|. \quad (26a)$$

Instead of the first two terms we may also write

$$-\left(\frac{|\tau + \tau'| + |\tau - \tau'|}{2} \right)^2.$$

Hence if we introduce the abbreviation

$$\tau^* = \frac{1}{2}(|\tau + \tau'| + |\tau - \tau'|) \quad (26b)$$

we get instead of (26a)

$$\Lambda = -\tau^{*2} - \tau^* = -\tau^*(\tau^* + 1) \quad (26c)$$

and eqn. (26) becomes

$$y(2-y)v'' + 2[1 + |\tau + \tau'| - y(\tau^* + 1)]v' + [\lambda - \tau^*(\tau^* + 1)]v = 0. \quad (27)$$

We already know from the equation of definition (25) that v behaves regularly at $y = 0$ and can therefore be expanded as a power series :

$$v = \sum \alpha_\nu y^\nu \quad (28)$$

The recurrence formula for the coefficients α_ν is obtained by substituting the series (28) in (27) and, for example, to equate the factor of y^ν to zero. In our case the recurrence formula consists of *two members* and runs :

$$\alpha_{\nu+1} \{ 2(\nu+1)\nu + 2(\nu+1)(1 + |\tau + \tau'|) \} + \alpha_\nu \{ -\nu(\nu-1) - 2\nu(\tau^* + 1) + \lambda - \tau^*(\tau^* + 1) \} = 0 \quad (29)$$

It is at this point that our polynomial method enters ; if we make the factor of α_ν vanish for a certain value, say $\nu = \mu$, then all the following coefficients α_ν for $\nu > \mu$ vanish. This not only secures convergence but also the regular behaviour of v at the second singular point $y = 2$ of the differential equation (corresponding to $x = 1$). Hence if by means of (25) and (18) we return from v to Θ and ψ , we have built up the desired proper function.

After the second line of (29) we force the recurrence to cease for $\nu = \mu$ by fixing the "proper value" λ as follows:

$$\lambda = \mu(\mu - 1) + 2\mu(\tau^* + 1) + \tau^*(\tau^* + 1).$$

This is equivalent to

$$\lambda = m(m + 1) \quad (30)$$

if we set

$$m = \mu + \tau^* \quad (31)$$

τ^* is, by (26b), a positive whole number, just like μ . For τ and τ' are integers (cf. above); further, $|\tau + \tau'|$ and $|\tau - \tau'|$ are positive integers, which are simultaneously either even or odd, so that half their sum is always an integer, and by (26b), τ^* is simply the greater of the two numbers $|\tau|$ and $|\tau'|$. Accordingly, by (31), m also becomes a *positive integer*.

It follows from (30), in consequence of (20), that

$$E = \frac{m(m + 1)h^2}{8\pi^2A} + \frac{\tau^2 h^2}{8\pi^2C} \left(\frac{1}{C} - \frac{1}{A} \right) \quad (32)$$

This is formula (7) of I, p. 443, already mentioned at the beginning of this section, except for unessential changes in the notation and with the essential difference that m^2 is replaced by $m(m + 1)$. Our present τ corresponds exactly to the earlier m_0 and, like m_0 , measures the "proper momentum" of the top (for τ belongs to the angle of rotation ϕ about the axis of the figure, that is, to the "proper rotation"). Also the *dropping out of the band lines on both sides of the zero line*, which was emphasised earlier in I, p. 445, is contained in our formulæ. For whereas μ , being the degree of our polynomial (at the same time the number of nodes in the θ direction), can assume all the values 0, 1, 2 . . ., we see from (31) that m is restricted to the values $m \geq \tau^*$. Finally, the *occurrence of a zero branch* (I, p. 446) can be explained on the theory of wave-mechanics. For if we start from our assumption (18), which corresponds to a motion compounded of precession and rotation, the selection rules for m differ from those which we have obtained in the case of pure rotation: *the transition $m \rightarrow m$ is not forbidden here*, as it was in the case of pure rotation [cf. § 6, eqn. (33)].

The proper functions of our problem are fully determined by the recurrence formula (29) [except for a multiplying constant, which follows from the normalising condition]. Concerning their analytical character we mention only that they are special cases of the hypergeometric series, namely, the so-called Jacobian polynomials. But it is characteristic of our "polynomial method" that it attains its goal without requiring any special knowledge of the theory of functions, and in the present case we do not need to have recourse to the full theory of the hypergeometric series.

§ 12. Concerning the Relationship Between Quantum Conditions in Wave-Mechanics and in the Older Quantum Theory

G. Wentzel * and L. Brillouin † have devised an instructive method of making the wave-mechanical methods of solution approximate to the processes of classical mechanics. We shall here give a form of this method which does not go so far as was originally intended by those authors. For we shall assume that the proper functions and proper values have been found wave-mechanically, and shall enlighten ourselves only on the relationship of the quantum conditions in the new and the old theory.

We revert to the beginning of this Chapter and, similarly to eqn. (6a), p. 4, we set

$$\psi = e^{\frac{2\pi i}{h} \int y dx} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

Here $\int y dx$ takes the place of the earlier S ; the amplitude-factor A , which was separate earlier, has been included in the exponent. By calling the variable of integration x , we indicate that for the present we are restricting ourselves to one-dimensional problems. The integral is to be thought of as an indefinite integral with an arbitrary but fixed lower limit x_0 and with the variable upper limit x .

The one-dimensional wave-equation for ψ is

$$\frac{d^2\psi}{dx^2} + \frac{8\pi^2m}{h^2}(E - V)\psi = 0$$

or, written more simply for what follows,

$$\frac{d^2\psi}{dx^2} + \left(\frac{2\pi}{h}\right)^2 p^2 \psi = 0 \quad . \quad . \quad . \quad . \quad (2)$$

Here the "momentum" p is defined classically according to the energy equation

$$\frac{p^2}{2m} + V = E$$

by

$$p = \sqrt{2m(E - V)} \quad . \quad . \quad . \quad . \quad (3)$$

The differential equation for ψ is of the *second order* and of the *first degree*. We show that it gives rise to a differential equation for y which, like Hamilton's equation for S , is of the *first order* and the *second degree*. For this purpose we form from (1)

$$\psi' = \frac{2\pi i}{h} y \psi, \quad \psi'' = \left[\frac{2\pi i}{h} y' - \left(\frac{2\pi}{h}\right)^2 y^2 \right] \psi \quad . \quad . \quad (3a)$$

* Zeitschr. f. Phys., **38**, 518 (1926). Besides the Kepler motion Wentzel also treats the Stark effect of the second order by this method in a particularly simple way. We shall derive his result in the next chapter, § 2, by a different method.

† Comptes Rendus, July, 1926.

and obtain from (2), by dividing by $\left(\frac{2\pi i}{h}\right)^2 \psi$,

$$\frac{h}{2\pi i} y' = p^2 - y^2 \quad . \quad . \quad . \quad . \quad . \quad (4)$$

This is the "Riccati equation" corresponding to eqn. (2).

Using $\frac{h}{2\pi i}$ as the argument of an expansion, we solve eqn. (4) by assuming

$$y = y_0 + \frac{h}{2\pi i} y_1 + \dots + \left(\frac{h}{2\pi i}\right)^{\nu} y^{\nu} + Y \quad . \quad . \quad . \quad (5)$$

To escape all difficulties of convergence we therefore break off the expansion at the ν^{th} member and call the remainder Y . By substituting (5) in (4) and comparing the form of the first powers of $\frac{h}{2\pi i}$, we find

$$\left(\frac{h}{2\pi i}\right)^0 \quad \left| \quad 0 = p^2 - y_0^2, \quad y_0 = p \quad . \quad . \quad . \quad . \quad (6)\right.$$

$$\left(\frac{h}{2\pi i}\right)^1 \quad \left| \quad y'_0 = -2y_0 y_1, \quad y_1 = -\frac{y'_0}{2y_0} \quad . \quad . \quad . \quad (7)\right.$$

$$\left(\frac{h}{2\pi i}\right)^2 \quad \left| \quad y'_1 = -y_1^2 - 2y_0 y_2, \quad y_2 = -\frac{y'_1 + y_1^2}{2y_0} \quad . \quad (8)\right.$$

We thus see that the successive approximations y_0, y_1, y_2, \dots are obtained without integration, simply by differentiations, from the approximation of zero order. This approximation y_0 itself corresponds to classical mechanics. For by the latter the function of action S is defined by

$$\frac{\partial S}{\partial q} = p \quad . \quad . \quad . \quad . \quad . \quad (9)$$

If, corresponding to our present notation, we here make $q = x$ and $S = \int y_0 dx$, (9) becomes identical with (6). Thus our series expansion (5) successively corrects the classical-mechanical solution, and makes it merge into our wave-mechanical solution.

Our next purpose is to bring the conditions of continuity of the new theory into relationship with the quantum conditions of the old theory. A glance at eqn. (1) suggests the following line of argument: ψ is to be one-valued and continuous along the real x -axis (or along that part of it which comes into question owing to the physical meaning of x). If we assume beyond this that ψ depends *analytically* on x , then the one-valuedness and continuity of ψ extends into the



FIG. 18.—Closed path for integration by the Wentzel-Brillouin method. The points $x = \pm a$ refer to the oscillator and give the extent of the classical orbit.

complex plane, that is, in a certain region in the neighbourhood of the real x -axis. In actual fact the nearest singular point (branch-point or point at infinity) will be at a finite distance from the x -axis. If in this part of the x -plane, namely in the shaded portion of Fig. 13, we describe a closed curve, starting from a point A and ending at a point B coincident with it, ψ must return to its initial value. But this means that the exponent of the right side of (1) can only vary, if at all, by an integral multiple n of $2\pi i$. So we postulate

$$\frac{2\pi i}{h} \left(\int_{x_0}^B y dx - \int_{x_0}^A y dx \right) = 2\pi i n,$$

or, written more simply,

$$\oint y dx = nh \quad . \quad . \quad . \quad . \quad (10)$$

in which we indicate the integration from A to B by \oint . Substituting in the expansion (5), we get

$$\oint y_0 dx + \frac{h}{2\pi i} \oint y_1 dx + \dots + \oint Y dx = nh \quad . \quad . \quad (11)$$

If we break off the expansion at the first term, we have, setting $y_0 = p$ and $x = q$, the quantum condition of the old theory

$$\oint p dq = nh.$$

The succeeding terms of the expansion therefore give those corrections which are necessary according to the new theory. In this process we assume that the expansion is carried so far that the remainder term Y makes no contribution, that is

$$\oint Y dx = 0 \quad . \quad . \quad . \quad . \quad (12)$$

We shall investigate only in special cases how the corrections are to be calculated and whether the condition (12) can be fulfilled.

We have yet to remark that on account of the one-valuedness of ψ the condition (10) can be imposed for any arbitrary closed path, but that we are justified in choosing paths for which the integration may easily be performed and does not become trivial. (Curves which enclose no singularities and therefore make $\oint y dx$ vanish do not interest us.) For the sequel we shall agree to carry out the integration around the whole region of the independent variable that comes into question physically, but excluding the end points of the region. First we shall consider the path of integration to embrace the x -axis closely on both sides as in Fig. 13; later we shall deform it so far as this is possible according to the regularity of the function to be integrated. The only singularities of y enclosed by our

path are the zero-points of ψ , since ψ and ψ' are not allowed to have singularities in the physical region. The number of zero points ["number of nodes" (*Knotenzahl*)] simultaneously determines the quantum number n in eqn. (10). It is easy to see that zero points of ψ can lie only in that part of the real axis where $p^2 > 0$. For by eqn. (2) it is only here that the curvature of the ψ -curve is concave to the x -axis, that is, it is only here that ψ is oscillatory in character. This part of the x -axis is identical with the extent of the orbit which was described according to the earlier theory. In view of this we may express our above convention about the path as follows: the path of integration is to be taken *round the region of the classical orbit*. In consequence of the fact that deformation of the path of integration is possible these two conventions amount to the same thing.

A. The Harmonic Oscillator

If m denotes the mass, ω_0 the classically calculated frequency of the free vibration of the oscillator and a the maximum elongation in the classical motion, we can set [cf. § 3, eqn. (1)]

$$V = \frac{m}{2} \omega_0^2 x^2, \quad E = \frac{m}{2} \omega_0^2 a^2. \quad (13)$$

Thus

$$p = \sqrt{2m(E - V)} = m\omega_0 \sqrt{a^2 - x^2}.$$

Hence by (6) and (7)

$$\left. \begin{aligned} y_0 &= m\omega_0 \sqrt{a^2 - x^2} \\ y_1 &= -\frac{1}{2} \frac{d \log y_0}{dx} = -\frac{1}{4} \left(\frac{1}{a+x} - \frac{1}{a-x} \right) \end{aligned} \right\} \quad (13a)$$

From y_0 and y_1 in (8) we easily calculate

$$y_2 = -\frac{1}{8m\omega_0} \frac{2a^2 + 3x^2}{(a^2 - x^2)^{3/2}}.$$

From this we see immediately that

$$\oint y_2 dx = 0 \quad (14)$$

so long as the path of integration encloses the two branch points $x = \pm a$; and this is so on account of our convention (cf. Fig. 13). Then the path of integration may be extended to infinity, and y_2 becomes proportional to $1/x^3$. With the same method of treatment we find that for $x \rightarrow \infty$

$$\oint y_1 dx \text{ merges into } -\frac{1}{2} \oint \frac{dx}{x},$$

where the last integral may be taken, for example, over a circle of sufficiently great radius. With the positive sense of traversing the path assumed in the figure it has the value $2\pi i$. Hence it follows that

$$\frac{h}{2\pi i} \oint y_1 dx = -\frac{h}{2} \quad (15)$$

To calculate $\oint y_0 dx$ we contract the path of integration most conveniently to the real axis from $-a$ to $+a$ and back, and set $x = a \sin \phi$. We get

$$\oint y_0 dx = m\omega_0 a^2 \int_0^{2\pi} \cos^2 \phi d\phi = m\omega_0 a^2 \pi.$$

By using the expression for E in (13), we may write

$$\oint y_0 dx = \frac{2\pi}{\omega_0} E = \frac{E}{\nu_0} \quad . \quad . \quad . \quad (16)$$

which is, of course, *au fond* identical with the earlier evaluation of the phase integral for the oscillator in I, Chap. IV, § 1, p. 197.

We can show that, just as the integral over y_2 vanishes, so all the subsequent integrals of our expansion (11) vanish. We prefer, however, to break off the expansion at the second member, that is, at y_1 and to take up the remaining members into the remainder Y . Anticipating that the remainder integral becomes zero, eqn. (11) becomes, with due regard to (15) and (16),

$$\frac{E}{\nu_0} - \frac{h}{2} = nh, \quad \text{that is,} \quad E = (n + \frac{1}{2})h\nu_0 \quad . \quad . \quad (17)$$

This is our familiar *wave-mechanical value for the energy of the oscillator*. If we had omitted the correction member in y_1 , we should have obtained *the value of the earlier theory* $E = nh\nu_0$.

We next determine the *remainder*. If we can show that Y (similarly to y_2 before) possesses for large values of x an expansion in negative powers of x , in which no term containing $1/x$ occurs, this simultaneously proves that the remainder integral vanishes, the same path of integration being chosen as previously. This is actually the case *if we use (1) to express y by means of ψ and substitute for ψ the proper function, free of singularities, determined from the wave-equation*.

By eqns. (8), (5a), and (2a) of § 3 we have

$$\psi = Ae^{-\xi^2/2} H_n(\xi), \quad \xi = \sqrt{\alpha}x, \quad \alpha = \frac{2\pi m\omega_0}{h} \quad . \quad . \quad (18)$$

where A = the normalising factor. Concerning $H_n(\xi)$ we need know only that it is a polynomial of the n^{th} degree; that is, it has the form $H_n(\xi) = c\xi^n + \dots$. From this it follows for great values of ξ that

$$\frac{H'_n(\xi)}{H_n(\xi)} = \frac{n}{\xi} + \dots$$

where the unwritten terms here and in the sequel, denote powers $\frac{1}{\xi^2}, \frac{1}{\xi^3}, \dots$.

By the equation of definition (1) and by (18)

$$y = \frac{h}{2\pi i} \frac{d \log \psi}{dx} = \frac{h \sqrt{\alpha}}{2\pi i} \frac{d \log \psi}{d\xi} = \frac{h \sqrt{\alpha}}{2\pi i} \left(-\xi + \frac{n}{\xi} + \dots \right) \quad (19)$$

To determine Y we start from eqn. (5) in which, as agreed, we make

$\nu = 1$. Multiplying by $\frac{2\pi i}{h\sqrt{\alpha}}$ we get, on account of (19),

$$\frac{2\pi i}{h\sqrt{\alpha}}Y = -\xi + \frac{n}{\xi} + \dots - \frac{2\pi i}{h\sqrt{\alpha}}y_0 - \frac{1}{\sqrt{\alpha}}y_1 \quad (20)$$

On account of (13a) we have for large values of x^*

$$y_0 = im\omega_0 x \left(1 - \frac{1}{2} \frac{a^2}{x^2} + \dots\right), \quad y_1 = -\frac{1}{2x} + \dots$$

In view of the definition of ξ and α in (18) it follows from this that

$$\left. \begin{aligned} -\frac{2\pi i}{h\sqrt{\alpha}}y_0 - \frac{1}{\sqrt{\alpha}}y_1 &= \frac{2\pi m\omega_0}{h\sqrt{\alpha}}x \left(1 - \frac{1}{2} \frac{a^2}{x^2} + \dots\right) \\ &+ \frac{1}{2\sqrt{\alpha}x} + \dots = \xi - \frac{1}{2} \frac{\alpha a^2}{\xi} + \frac{1}{2\xi} \end{aligned} \right\} \quad (21)$$

In this we have yet to express the quantity a^2 in terms of the energy value E of eqns. (13) and (17). On account of

$$E = \frac{m}{2} \omega_0^2 a^2 = \left(n + \frac{1}{2}\right) h \frac{\omega_0}{2\pi} \quad (21a)$$

we have

$$\frac{\pi m \omega_0}{h} a^2 = n + \frac{1}{2}$$

for which we may also write

$$\frac{\alpha a^2}{2} = n + \frac{1}{2}$$

The right-hand side of (21) thus reduces to

$$\xi - \frac{n + \frac{1}{2}}{\xi} + \frac{1}{2\xi} + \dots = \xi - \frac{n}{\xi} + \dots,$$

and this exactly cancels out with the first two members of the expansion (20). Thus our remainder Y has in fact for large values of ξ the required property of vanishing more rapidly than $1/\xi$, so that eqn. (12) is proved for our remainder integral.

We must point out that this proof, besides assuming the existence of the proper function, also uses [cf. (21 a)] the wave-mechanical expression of the proper value E . We cannot therefore claim to have deduced this value in eqn. (17) according to the method of Wentzel and Brillouin. But we see clearly in our example how the quantum rules of the old theory are connected with the system of wave-mechanics.

B. The Kepler Problem

We are here interested only in the radial part R of the proper function, which by (3) and (3a) of § 7 satisfies the differential equation

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \left\{ \left(\frac{2\pi}{h}\right)^2 p^2 - \frac{l(l+1)}{r^2} \right\} R = 0 \quad (22)$$

* The sign of y_0 is determined by the fact that in the path of integration of Fig. 18 we must take the square root of y_0 as *positive* for increasing real values of x .

where

$$p^2 = 2m(E - V) = 2m\left(E + \frac{Ze^2}{r}\right) \quad (22a)$$

We introduce the abbreviations

$$A = 2mE, \quad B = mZe^2, \quad C = - (l + 1)^2 \left(\frac{h}{2\pi}\right)^2 \quad (23)$$

which differ slightly from those defined in eqn. (3a), § 7, but agree exactly with the abbreviations A, B, C, of the older theory in I, Note 8, p. 570. In particular, it must be noticed that our present quantity $l + 1$ is identical with the quantum number k there used. From the definition of C we get

$$-l(l + 1) = - (l + 1)^2 + (l + 1) = \left(\frac{2\pi}{h}\right)^2 C + (l + 1).$$

If we make all the substitutions in (22), we get

$$\frac{d^2 R}{dr^2} + \frac{2}{r} \frac{dR}{dr} + \frac{l + 1}{r^2} R + \left(\frac{2\pi}{h}\right)^2 \left(A + \frac{2B}{r} + \frac{C}{r^2}\right) R = 0 \quad (24)$$

As in eqn. (1), we now insert y in place of R by means of the assumption

$$R = e^{\frac{2\pi i}{h} \int y dr} \quad (24a)$$

and obtain from (24) an equation analogous to Riccati's eqn. (4) [cf. also (3a)]

$$\frac{h}{2\pi i} \left(\frac{dy}{dr} + \frac{2y}{r} + \frac{\lambda}{r^2} \right) = P^2 - y^2 \quad (25)$$

with the abbreviations

$$P^2 = A + \frac{2B}{r} + \frac{C}{r^2}, \quad \lambda = \frac{h}{2\pi i} (l + 1).$$

As in (5) we then expand y in terms of the parameter $h/2\pi i$. (The fact that the quantity λ itself still contains the factor $h/2\pi i$ need not disturb us, since we can group the terms and break off the expansion at will.) So we write

$$y = y_0 + \frac{h}{2\pi i} y_1 + \dots + Y \quad (26)$$

and, exactly as in (6) and (7), by substitution in (25) we get

$$y_0 = \sqrt{A + \frac{2B}{r} + \frac{C}{r^2}}, \quad y_1 = -\frac{1}{2y_0} \left(\frac{dy_0}{dr} + \frac{2y_0}{r} + \frac{\lambda}{r^2} \right) \quad (27)$$

In the Kepler problem the range of the co-ordinates extends from $r = 0$ to $r = \infty$. The two roots of $P = 0$, which we earlier called r_{max} and r_{min} , lie on the real axis (cf. I, Fig. 124, p. 551). By our convention on p. 135, the path of integration runs round these two points and afterwards goes to infinity in the r -plane. The point $r = 0$ must be considered ultimately as a possible singularity.

The postulate that the proper function must be one-valued requires that $\oint y dr$ must equal h times an integer, which we shall now call n_r (radial quantum number). From eqn. (26) we therefore get the equation, analogous to (11),

$$\int y_0 dr + \frac{h}{2\pi i} \oint y_1 dr + \dots + \oint Y dr = n_r h \quad (28)$$

We shall show that here even the *first* correction term disappears, from which the familiar fact follows that *the energy-levels of the Kepler problem found in the older theory remain preserved unchanged in wave-mechanics*. For this purpose we consider the power expansions of y_0 and y_1 in the neighbourhood of the points $r = 0$ and $r = \infty$.

For $r = 0$ we have

$$y_0 = \frac{\sqrt{C}}{r} \left(1 + \frac{B}{C} r + \dots \right), \quad \frac{dy_0}{dr} = -\frac{\sqrt{C}}{r^2} + \dots,$$

$$y_1 = -\frac{r}{2\sqrt{C}} \left(-\frac{\sqrt{C}}{r^2} + \frac{2\sqrt{C}}{r^2} + \frac{\lambda}{r^2} + \dots \right) \quad (29)$$

By (23)

$$\sqrt{C} = \frac{h}{2\pi i} (l + 1) \quad (29a)$$

that is, equals our value of λ : the fact that the signs of these two quantities \sqrt{C} and λ also agree is evident from the concluding remark in I, Note 6, c, p. 552. Hence in the neighbourhood of $r = 0$, we have, by (29) and (29a),

$$y_1 = -\frac{1}{r} + \dots \quad (29b)$$

On the other hand, we have for $r = \infty$ (except for terms that vanish to a higher order for $r = \infty$)

$$y_0 = \sqrt{A} \left(1 + \frac{B}{A} \frac{1}{r} + \dots \right), \quad \frac{dy_0}{dr} = -\frac{B}{\sqrt{A}} \frac{1}{r^2} + \dots$$

$$y_1 = -\frac{1}{2\sqrt{A}} \left(\dots + \frac{2\sqrt{A}}{r} + \dots \right) = -\frac{1}{r} + \dots \quad (29c)$$

If we now calculate $\oint y_1 dr$ in accordance with I, Fig. 124, p. 551, the two points $r = 0$ and $r = \infty$ give *opposite* contributions, namely $+2\pi i$ and $-2\pi i$ respectively. In all, then, we have, as was asserted,

$$\oint y_1 dr = 0. \quad (30)$$

The same result would be obtained if we were to calculate the subsequent correction terms in the expansion (28). Assuming that the

remainder integral vanishes, eqn. (28) therefore gives us the quantum condition

$$\oint y_0 dr = \oint \sqrt{A + 2\frac{B}{r} + \frac{C}{r^2}} dr = 2\pi i \left(\frac{B}{\sqrt{A}} - \sqrt{C} \right) = n_r h \quad (31)$$

just as in the earlier theory, I, Note 8, p. 569.

In dealing with the remainder term Y we shall now find it convenient to have the term in y_1 included in it, that is, to write instead of (26)

$$Y = y - y_0 \quad (32)$$

By (24a) and (27),

$$y = \frac{h}{2\pi i} \frac{d \log R}{dr}, \quad y_0 = \sqrt{A + 2\frac{B}{r} + \frac{C}{r^2}}$$

The expression for R , by eqn. (13), § 7, runs

$$R = e^{-\rho/2} \rho^l L(\rho) \quad (33)$$

if we omit a normalising factor which is of no account here. L is a polynomial connected with that of Laguerre. All we need know about it is that it is of degree n_r ; ρ is proportional to r .

As in eqn. (18), we get from (33) that

$$\frac{d \log R}{d\rho} = -\frac{1}{2} + \frac{l + n_r}{\rho} + \dots$$

The relation between ρ and r is, by eqns. (3a) and (4a) of § 7, if we express it in terms of our present quantity A , eqn. (23),

$$\rho = \frac{4\pi i}{h} \sqrt{A} r \quad (33a)$$

Consequently we get

$$\frac{h}{2\pi i} \frac{d \log R}{dr} = 2\sqrt{A} \left(-\frac{1}{2} + \frac{l + n_r}{\rho} \right) + \dots \quad (34)$$

For large values of r the expansion of y_0 (cf. above) is

$$y_0 = \sqrt{A} \left(1 + \frac{B}{A} \frac{1}{r} + \dots \right) = \sqrt{A} \left(1 + \frac{B}{\sqrt{A}} \frac{4\pi i}{h} \frac{1}{\rho} + \dots \right) \quad (35)$$

If we substitute (34) and (35) in (32) we get as the expression for Y for great values of ρ

$$Y = 2\sqrt{A} \left(-\frac{1}{2} + \frac{l + n_r}{\rho} + \dots \right) - \sqrt{A} \left(1 + \frac{B}{\sqrt{A}} \frac{4\pi i}{h} \frac{1}{\rho} + \dots \right).$$

If we take the integral $\int Y dr$ around the point $r = \infty$, it follows from this that

$$\int Y dr = \frac{h}{2\pi i} \frac{1}{2\sqrt{A}} \int Y d\rho = h(l + n_r) - 2\pi i \frac{B}{\sqrt{A}}.$$

On the other hand, we get for the expansion of y and y_0 at $r = 0$ (observing that the polynomial L in (33) becomes equal to a non-vanishing constant for $\rho = 0$, by eqn. (29), § 7, and likewise its first derivative):

$$y = 2\sqrt{A}\left(\dots + \frac{l}{\rho} + \dots\right), \quad y_0 = \frac{\sqrt{C} + \dots}{r} = \frac{4\pi i}{h} \frac{\sqrt{AC}}{\rho} + \dots$$

This point contributes the following amount to our integral:

$$\int Y dr = \frac{h}{2\pi i} \frac{1}{2\sqrt{A}} \int Y d\rho = -lh + 2\pi i \sqrt{C}.$$

By adding together the contributions for $r = 0$ and $r = \infty$ we get for the integration along the whole path

$$n_r h - 2\pi i \left(\frac{B}{\sqrt{A}} - \sqrt{C} \right),$$

that is, if we take into consideration the quantum condition (31), the result *nil*.

It is true that here, as in the case of the oscillator, we have assumed the correct quantum condition, eqn. (31) in this instance. That is, we have not really derived this condition but again have only linked up the previous method of quantising with wave-mechanics.

We call special attention to this here because the argument on this point in the German edition moves in a circle: the vanishing of the residual term was derived from the quantum condition and the quantum condition was derived from the vanishing of the residual term.

CHAPTER II

PERTURBATIONS. DIFFRACTION PROBLEMS. THE SPINNING ELECTRON

§ 1. Schrödinger's Theory of Perturbations

AMONG the many beautiful results of wave-mechanics the theory of perturbations developed by Schrödinger* occupies a special position. It is much simpler and more lucid than the astronomical theory of perturbations of classical mechanics, and need not even fear the complication of the three-bodies problem (helium problem, § 8).

A. The Perturbation Scheme for the Non-Degenerate Case

Following Schrödinger closely, we consider a problem which we can solve wave-mechanically by omitting the perturbation terms. Let the proper functions of the non-perturbation problem be ψ_k , and the corresponding proper values E_k . For the present these will be considered as *simple*. We write the wave-equation of the unperturbed problem as

$$L\psi + E_p\psi = 0 \quad . \quad . \quad . \quad . \quad (1)$$

As in Chap. I, § 9 C, L denotes a linear partial differential expression of the second order. We may assume that L is "self-adjoint" [cf. Chap. I, § 11, A, eqn. (10)]. We include the factor p (weight-function) because we must in general use curvilinear co-ordinates q suitable to the problem in question, for which the functional determinant D becomes added as a factor to E [cf. likewise Chap. I, § 11, eqn. (10)]. Instead of D we choose the more general term p , because in certain cases (cf. the Stark effect in § 2) we shall also include other factors in p .

The condition of orthogonality for two proper functions ψ_i, ψ_k follow from the general theorem of Green, eqn. (20), § 9, Chap. I; the right-hand side of the eqn. (20) vanishes on account of the boundary condition to be imposed on the proper functions. In consequence of the assumption that L is self-adjoint we have $M = L$:

$$\int \{\psi_i L(\psi_k) - \psi_k L(\psi_i)\} dq = 0 \quad . \quad . \quad . \quad (1a)$$

dq denotes the product of the differentials of the co-ordinates q ; on account of (1), it follows that

$$(E_i - E_k) \int \psi_i \psi_k p dq = 0.$$

* Cf. for this whole section "Quantisation as a Problem of Proper Values," Part III, *Collected Papers*, p. 62.

Hence the condition of orthogonality ($i \neq k$) is

$$\int \psi_i \psi_k p dq = 0 \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The normalising condition is

$$\int \psi_k^2 p dq = 1 \quad . \quad . \quad . \quad . \quad (3)$$

When a perturbation is superimposed, the expression L becomes changed by a small term, whose smallness we suppose measured by a parameter λ . If the perturbation occurs in the potential energy V the perturbation term is multiplied by ψ . With this assumption we write

$$L(\psi) + \text{Exp}\psi = \lambda s\psi \quad . \quad . \quad . \quad (4)$$

and take s as standing for a function of the co-ordinates, given by the type of perturbation.

Starting from the k^{th} proper state, we set

$$\mathbf{E} = \mathbf{E}_k + \lambda \epsilon, \quad \psi = \psi_k + \lambda \phi \quad . \quad . \quad . \quad (4a)$$

and obtain from (4), neglecting the members in λ^2 ,

$$L(\psi_k) + \lambda L(\phi) + E_k r \psi_k + \lambda E_k r \phi + \lambda \epsilon r \psi_k = \lambda s \psi_k.$$

On account of (1) the sum of the first and third members vanishes, and after cancelling out the common factor λ

$$L(\phi) + E_k p \phi = (s - \epsilon p) \psi_k \quad . \quad . \quad . \quad . \quad (5)$$

The left-hand side of this *non-homogeneous* equation is of the same form as that of the *homogeneous* eqn. (1). On the right-hand side the quantity ϵ is at present unknown, but can be determined from the true theory of perturbations by the general theorem: *if a non-homogeneous equation of the form (5) is to be soluble at all, that is, if it is to have a continuous solution, the right-hand side must be "orthogonal" to the solution of the homogeneous equation, that is, to the proper function ψ_k .*

To prove this we have only to multiply both sides of eqn. (5) by ψ_k and to integrate over the region of co-ordinates. As in eqn. (1a) the left-hand side becomes transformed into

$$\int \phi(L(\psi_k) + E_k p \psi_k) dq$$

and vanishes on account of the differential equation (1). Thus the right-hand side of (5) gives

$$\int (s - \epsilon p) \psi_k^2 dq = 0,$$

which was to be proved.

This theorem is a corner-stone of the theory of integral equations; its general importance, however, was recognised earlier by Lord Rayleigh in his Theory of Sound. It becomes evident if we explain it by using the vibrating membrane as an example. The differential equation for its periodic proper vibrations is

$$\Delta u + k^2 u = 0, \quad k^2 = \frac{\rho \omega^2}{S} \quad (6)$$

where ρ denotes the density per unit surface, ω the circular frequency of

the proper vibration, and S the tension acting in the membrane. If we now allow an external transversal pressure $P(x, y)$ to act which may be distributed anyhow over the membrane and whose period ω coincides with a proper vibration of the membrane, then it will excite the membrane to continually increasing vibrations. The periodic final state would be a "resonance catastrophe." If this is not to occur, P cannot on the whole perform work on the vibrating membrane. The pressure may therefore act, for example, only in the lines of nodes, or must have the same magnitude in oppositely vibrating sectors, and so forth. Since $Pd\sigma$ is the force and u the distance traversed from the position of rest, the general condition for the performance of no work by P is

$$\int P u d\sigma = 0 \quad . \quad . \quad . \quad . \quad . \quad (7)$$

Applying this to our case we have $P = (s - \epsilon p) \psi_k$, $u = \psi_k$, and $d\sigma = dq$. Eqn. (7) thus demands that

$$\int (s - \epsilon p) \psi_k^2 dq = 0 \quad . \quad . \quad . \quad . \quad . \quad (7a)$$

This is what was meant above when we postulated that the right-hand side of (5) must be "orthogonal" to the proper function ψ_k .

From (7a) it follows at once that

$$\epsilon = \frac{\int s \psi_k^2 dq}{\int p \psi_k^2 dq} \quad . \quad . \quad . \quad . \quad . \quad (8)$$

or, if ψ_k is normalised in accordance with (3), still more simply

$$\epsilon = \int s \psi_k^2 dq \quad . \quad . \quad . \quad . \quad . \quad (8a)$$

The value of ϵ determined in this way is now substituted on the right-hand side of (5), and we expand the expression so obtained in terms of the proper functions ψ_i of the non-perturbation problem, having previously divided by p .

That is, let

$$\left(\frac{s}{p} - \epsilon\right) \psi_k = \sum A_i \psi_i \quad . \quad . \quad . \quad . \quad . \quad (9)$$

We expand the required function ϕ in the same way,

$$\phi = \sum B_i \psi_i \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Since, by (1),

$$L(\psi_i) = -E_i p \psi_i$$

and hence also

$$-L(\phi) = -\sum B_i E_i p \psi_i$$

eqn. (5) now runs

$$\sum_i B_i (E_k - E_i) p \psi_i = p \sum_i A_i \psi_i.$$

By equating the coefficients we see that

$$B_i = \frac{A_i}{E_k - E_i} \quad (10a)$$

and hence, on account of (10),

$$\phi = \sum^{(i)} \frac{A_i \psi_i}{E_k - E_i} \quad (11)$$

A_i may be calculated in Fourier's manner from the equation of definition (9); taking into account (2) and (3),

$$A_i = \int (s - \epsilon p) \psi_k \psi_i dq = \int s \psi_k \psi_i dq \quad (12)$$

and in particular, if we consider (7a), $A_k = 0$. Let the dropping out of the term with k from the sum (11) be denoted as usual by Σ' .

This solves our perturbation problem in an extremely clear and general way: substituting (8a), (11) and (12) in (4a) gives for the perturbed proper value and the perturbed proper function

$$E = E_k + \lambda \int s \psi_k^2 dq, \quad \psi = \psi_k + \lambda \sum_i' \frac{\psi_i \int s \psi_k \psi_i dq}{E_k - E_i} \quad (13)$$

We have yet to say a word about the possibility of the expansions (9) and (10). This is well known from special cases (trigonometric functions and spherical harmonics) and is proved in mathematical books for arbitrary proper functions—most rigorously in the theory of integral equations, the function to be expanded sometimes being subject to restrictions. Here we shall only emphasise that the sequence of proper functions must be complete; in the case of a Fourier series, for example, no cosine member may be omitted if the series is to represent a general function. We shall, however, not enter into the Completeness Relation (*Vollständigkeits Relation* *) set up to test this condition, as it is difficult to manipulate in an individual case, and shall restrict ourselves to making a few remarks that will be useful in the sequel. In many problems of wave-mechanics (Kepler problem, oscillating rotator) there is, as we know, not only a discrete but also a continuous spectrum of proper values. It would be inadmissible to leave out of the expansion the proper functions corresponding to this continuous spectrum. Thus in such cases we have to regard the summation in (13) as including also an integral over the continuous proper values. For later use we indicate this in (13a) by using E' to denote the continuous energy parameter and $\psi(E')dE'$ the corresponding proper function

$$\psi = \psi_k + \lambda \sum_i' \frac{\psi_i \int s \psi_k \psi_i dq}{E_k - E_i} + \lambda \int \frac{\psi(E') \int s \psi_k \psi(E') dq}{E_k - E'} dE' \quad (13a)$$

* Courant-Hilbert, Chap. II, § 1, eqns. (9) and (9'). We shall derive this relation in § 7, B.

B. *Perturbations in the Case of Several Proper Values*

The method so far described is in itself sufficient to deal with some of the most important perturbation problems, such as the Stark effect, § 2, the theory of dispersion, § 3, the photo-electric effect, § 4. The full beauty of Schrödinger's theory of perturbations, however, manifests itself only under somewhat more complicated conditions, namely when the problem is degenerate. The most striking example for the more refined theory of perturbation then called into action is the helium problem, § 8.

Starting from eqn. (1) we assume that to the proper value E_k there belong several linearly independent * proper functions

$$\psi_{k1}, \psi_{k2}, \dots, \psi_{ki}, \dots, \psi_{ka}.$$

We then speak, as in Chap. I, § 3, C, of an $(a - 1)$ -fold degeneracy. We shall assume that these ψ_{ki} 's are normalised to 1, and have been made orthogonal not only towards the other proper functions ψ but also among themselves. This does not yet, however, define them uniquely. Rather, we can subject them to any arbitrary *orthogonal transformation* without disturbing the character of their normalisation and orthogonality. For if we set

$$\psi_{kh} = \sum_1^{\alpha} \beta_{hi} \psi_{ki} \quad . \quad . \quad . \quad (14)$$

where the β 's stand for the coefficients of any arbitrary orthogonal substitution in space of α dimensions,

$$\text{and likewise} \quad \psi_{kh'} = \sum_1^{\alpha} \beta_{h'i} \psi_{ki} \quad . \quad . \quad . \quad (14a)$$

and calculate the characteristic integrals of eqns. (2) and (3) with these ψ_{kh} 's and $\psi_{kh'}$'s we find, on account of the orthogonality of the ψ_{ki} 's on the one hand and the β 's on the other,

$$\begin{aligned} \int \psi_{kh} \psi_{kh'} p \, dq &= \sum_i \sum_j \beta_{hi} \beta_{h'j} \int \psi_{ki} \psi_{kj} p \, dq \\ &= \sum_i \beta_{hi} \beta_{h'i} = \delta_{hh'} = \begin{cases} 0 & \dots h \neq h', \\ 1 & \dots h = h'. \end{cases} \end{aligned}$$

This means: the equations (2) and (3) are also satisfied for the ψ_{kh} 's if they were satisfied for the ψ_{ki} 's. Schrödinger's theory of perturbations, as we shall see, makes use of this arbitrariness in the choice of the proper functions in an extremely elegant way.

* "Linearly independent" obviously signifies in our case that no relation of the form

$$\sum c_i \psi_{ki} = 0$$

holds between the ψ_{ki} 's, where the c_i 's are constant and the index i assumes any of the values of the numbers 1 . . . α .

The existence of multiple proper values is known from classical mechanics, particularly from the theory of the vibrating membrane. In the sequel we shall speak of the vibrating *membrane* and not of the vibrating plate (Chladni's figures)—which is better known experimentally—because the membrane is simpler mathematically and is more closely related to the problems of wave-mechanics. Let us consider as the simplest case a *rectangular membrane* with fixed edges $x = 0$ and a , $y = 0$ and b . The solution of eqn. (6) is then (except for an arbitrary amplitude factor)

$$u = \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{b}.$$

The corresponding proper value is, by (6),

$$k_{m,n}^2 = \pi^2 \left(\frac{m^2}{a^2} + \frac{n^2}{b^2} \right).$$

It is *simple*, if a and b are incommensurable, for then there are no two numbers m and n that lead to the same value of k .

The position is different in the case of the *square membrane* $b = a$ (or, more generally, in the case of a rectangular membrane with commensurable edges). Then

$$k_{m,n}^2 = \frac{\pi^2}{a^2} (m^2 + n^2) \quad . \quad . \quad . \quad (15)$$

Interchanging m and n leaves the proper value unchanged, but alters the geometrical character of the proper functions. In actual fact the two states of vibration

$$\left. \begin{aligned} u_1 &= \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \\ u_2 &= \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{a} \end{aligned} \right\} \quad (16)$$

in general differ from one another. For example, u_1 has $m + 1$ nodal lines in the x -direction, u_2 has $n + 1$, and so forth. Only in the case $m = n$ do u_1 and u_2 become equal.

With the exception of the fundamental vibration $m = n = 1$, and the over-tones $m = n$ that are harmonic* to it, the proper values of the square membrane are at least two-fold. A higher degree of degeneracy occurs under certain conditions connected with the theory of numbers.†

In the case of two-fold proper values all functions of the family

$$v = \lambda_1 u_1 + \lambda_2 u_2 \quad . \quad . \quad . \quad (17)$$

(where λ_1 and λ_2 are arbitrary) become proper functions simultaneously with (16). If we wish to ensure that the v 's are normalised just as the u 's, we write instead

$$\left. \begin{aligned} v_1 &= \cos \gamma u_1 + \sin \gamma u_2, \\ v_2 &= -\sin \gamma u_1 + \cos \gamma u_2, \end{aligned} \right\} \quad (17a)$$

* If $m = n$, it follows from (6) and (15) that $\omega_{n,n} = n\omega_{1,1}$, that is, $\omega_{n,n}$ is a harmonic overtone of the fundamental $\omega_{1,1}$; on the other hand, $\omega_{n,m}$ is in general non-harmonic.

† Cf. the instructive book by F. Pockels (based on lectures by F. Klein): "Über die partielle Differentialgleichung $\Delta u + k^2 u = 0$, Teubner, 1891; in particular, pp. 79 et seq. The degree of the degeneracy is related to the number of ways in which the number $m^2 + n^2$ can be divided into primes.

analogously to (14). Fig. 14 shows how strongly the geometrical character of the proper vibration varies with λ_1 , λ_2 (respectively, γ). Still more complex manifolds* of nodal lines occur for $n > 2$. All types of vibration in the figure are equivalent among themselves. The cases $\lambda_2 = 0$, $v \sim u_1$, and $\lambda_1 = 0$, $v \sim u_2$, in which the nodal lines are straight and parallel to the sides of the square, are in no wise favoured above the other types of vibrations.

Now if the state of vibration is a little perturbed, there is no reason whatsoever why the perturbed state should follow continuously on the functions u_1 or u_2 . Rather, it will be developed, according to the nature of the perturbation, from that state of vibration of the family which is most suited to the acting perturbation. Suppose we act on the membrane, for example, by pressing lightly on it with a finger [special case of the pressure P in eqn. (7)]. Then that proper vibration of the family will be favoured, whose nodal lines runs through the point at which the pressure is exerted.

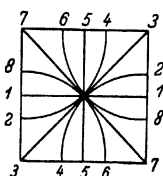


FIG. 14.

Nodal lines (Knotenlinien) of the square membrane
 $v = \lambda_1 u_1 + \lambda_2 u_2$.

$$\left. \begin{aligned} u_1 &= \sin \frac{m\pi x}{a} \sin \frac{n\pi y}{a} \\ u_2 &= \sin \frac{n\pi x}{a} \sin \frac{m\pi y}{a} \end{aligned} \right\} m = 1, \quad n = 2.$$

The number attached to the curves have the following meaning:

Curve	1	2	3	4	5	6	7	8
λ_1	1	1	1	$\frac{1}{2}$	0	$\frac{1}{2}$	1	1
λ_2	0	$-\frac{1}{2}$	-1	-1	1	1	1	$\frac{1}{2}$

The state of affairs is the same for the *circular membrane* represented in Fig. 15. Its proper vibrations are represented by

$$u = J_m(k_m, n \, r) \frac{\cos}{\sin} m\phi \quad . \quad . \quad . \quad (18)$$

The proof follows from the calculations of § 2, B, Chap. I, if we transpose it from three to two-dimensions, that is, from spherical harmonics to trigonometrical functions. J_m is the Bessel function of the integral index m , k_m, n the n th root of the equation $J_m(ka) = 0$, a the radius of the membrane, whose circumference we assume fixed. k_m, n is at the same time the general proper value of the membrane. An inspection of (18) shows that it is double. The only simple states are those ($m = 0$) whose nodal lines are purely circular

$$u = J_0(k_0, n \, r), \quad . \quad . \quad . \quad (18a)$$

to which the fundamental state $m = n = 0$ also belongs:

$$u = J_0(k_0, 0 \, r) \quad . \quad . \quad . \quad (18b)$$

* Cf. Pockels, loc. cit., p. 80.

Proper values higher than double do not occur in the case of the circular membrane. We shall find an exactly similar state of affairs in the helium problem: a simple ground state (S-state) and double excited states.

The linear family of proper functions which, belonging to the same $k_{m,n}$ with $m > 0$, interpolates itself between the two proper functions combined in (18), is represented by

$$v = \cos \gamma u_1 + \sin \gamma u_2 = J_m(k_{m,n} r) \cos m(\phi - \gamma).$$

Whereas the circular lines of nodes lie fixed within the family, the straight radial nodal lines are variable with γ . A superimposed perturbation selects from them that position γ , which is suited to the nature of the perturbation.

Fig. 15 represents the case $m = 1$, $n = 2$: one diameter of undetermined position is a nodal line. The circular nodal lines are reduced to the edge of the membrane and to a circle whose radius corresponds with the first root of $J_1(ka) = 0$.

We now return to wave-mechanics, and consider eqn. (4) with the perturbation function s and assume E_k to be an α -fold proper value. The assumption (4a) now also yields eqn. (5) but with the difference that here any of the α proper functions ψ_{ki} can be substituted for ψ_k . Meanwhile we know from the example of the membrane that these ψ_k 's play no favoured part, but that rather the perturbation chooses from the α -fold family (14) those proper functions to which they can attach themselves continuously. So in the right-hand side of (5) we write in place of a special ψ_k the general ψ_{kh} and keep open the choice of the substitution coefficients β_{ki} contained in it. How are these coefficients to be chosen? The answer is given by the theorem enunciated under A, which is expressed as follows for the present case: the right-hand side of the non-homogeneous equation must be orthogonal to *all* the solutions of the homogeneous equation. In place of the *one* condition (7a) this gives the α conditions

$$\int (s - \epsilon p) \psi_{kh} \psi_{kj} d\gamma = 0, \quad j = 1, 2 \dots \alpha.$$

By substituting the expression (14) for ψ_{kh} , we get (omitting the index h in β for the present)

$$\sum_i \beta_i \int (s - \epsilon p) \psi_{ki} \psi_{kj} dq = 0 \quad . \quad . \quad . \quad (19)$$

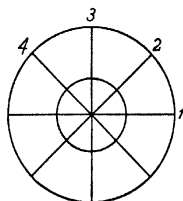


FIG. 15.
Lines of nodes of the circular membrane
 $v = J_m(k_{m,n} r) \cos m(\phi - \gamma)$,
 $m = 1, n = 2$.

The key to the numbers is

Curve	1	2	3	4
γ	$-\pi/2$	$-\pi/4$	0	$+\pi/4$

in which, by (23), we have for $l \neq k$

$$A_{ki} = \int \psi_{kh} \psi_{ki} dq.$$

The final result is, analogously to eqn. (13),

$$E = E_k + \lambda \epsilon_k$$

$$\psi = \psi_{kh} + \lambda \phi_k = \psi_{kh} + \lambda \sum_i' \sum_h \frac{A_{hi} \psi_{hi}}{E_k - E_i} \quad (26)$$

The dash attached to Σ_i' denotes as in (13) that the value $l = k$ is to be omitted in the summation. The index h numbers the α branches into which the degenerate problem splits up on account of the perturbation and corresponds to the α solutions of our algebraic eqn. (22).

Strictly speaking, all this is only a first step to the complete solution of the problem of perturbations, which would have to be supplemented by the approximations of the second, third . . . order, that is, by terms involving λ^2, λ^3 As in all calculations of perturbations these higher degrees of approximation become rather unmanageable in general, in contrast with the first approximation which, as we have seen, has been worked, for the degenerate as well as for the non-degenerate case, into an extremely elegant scheme by Schrödinger. In the Stark effect we shall also calculate the second degree of approximation without excessive difficulty.

§ 2. The Stark Effect

In the equation of the Kepler problem, eqn. (1) § 7, Chap. I, we add to the potential energy $V = -Ze^2/r$ the term eFx , as a homogeneous and relatively small external field F is assumed to act in the x -direction. We set

$$\frac{8\pi^2\mu e}{h^2} F = x$$

(we have called the electronic mass μ ; x replaces the parameter λ of the preceding section) and get in place of the eqn. (1) just mentioned,

$$\Delta\psi + \frac{8\pi^2\mu}{h^2} \left(E + \frac{Ze^2}{r} \right) \psi = x\psi \quad (1)$$

The literal application of the procedure described in § 1 would consist in expanding the right-hand side of (1) in terms of the proper functions of the Kepler problem

$$\psi = R(r)P_l^m(\cos\theta)e^{im\phi}.$$

This method would exactly correspond to Bohr's treatment of the Stark effect in *Atombau*, Chap. V, § 6 (cf. I, Chap. V, §§ 4 and 5) and has been carried out by Schrödinger in § 5 of his third communication on "Quantisation as a Problem in Proper Values." It is more illuminating and mathematically more satisfactory, however, to seek the method analogous

154 Chapter II. Perturbations. Diffraction Problems

to that of Schwarzschild and Epstein, that is, to postpone the introduction of perturbations and first to separate the problem rigorously in parabolic co-ordinates. Cf. §§ 3 and 4 of Schrödinger's third communication.

A. The Wave-equation in Parabolic Co-ordinates

The parabolic co-ordinates* ξ, η, ϕ are to be defined exactly as in I, Note 11, p. 587, eqns. (1) to (8):

$$x = \frac{1}{2}(\xi^2 - \eta^2), \quad y = \xi\eta \cos \phi, \quad z = \xi\eta \sin \phi \quad (2)$$

The kinetic energy, expressed as a function of the co-ordinates and momenta, is, by eqn. (10) of the same Note:

$$T(q, p) = \frac{1}{2\mu(\xi^2 + \eta^2)} \left\{ p_\xi^2 + p_\eta^2 + \left(\frac{1}{\xi^2} + \frac{1}{\eta^2} \right) p_\phi^2 \right\}.$$

If we adopt the notation of eqn. (1), § 11, Chap. I, for our present co-ordinates, we have

$$\begin{aligned} [\xi, \eta] &= [\eta\eta] = \frac{1}{2\mu(\xi^2 + \eta^2)}, \quad [\phi\phi] = \frac{1}{2\mu\xi^2\eta^2} \\ [\xi\eta] &= [\eta\phi] = [\phi\xi] = 0 \end{aligned}$$

and by (14) and (15) of the same section

$$(2\mu)^{-3/2} D = (\xi^2 + \eta^2) \xi \eta.$$

From the eqn. (10) of the Note it follows that the *wave-equation of the Stark-effect* is

$$\left. \begin{aligned} & \frac{\partial}{\partial \xi} \left(\xi \eta \frac{\partial \psi}{\partial \xi} \right) + \frac{\partial}{\partial \eta} \left(\xi \eta \frac{\partial \psi}{\partial \eta} \right) + \frac{\partial}{\partial \phi} \left(\frac{\xi^2 + \eta^2}{\xi \eta} \frac{\partial \psi}{\partial \phi} \right) \\ & + \frac{8\pi^2\mu}{h^2} \xi \eta (\xi^2 + \eta^2) \left\{ E + \frac{2Ze^2}{\xi^2 + \eta^2} - \frac{eF}{2} (\xi^2 - \eta^2) \right\} \psi = 0 \end{aligned} \right\} \quad (3)$$

In this we have set [cf. (2)]

$$V = -\frac{Ze^2}{r} + eFx = -\frac{2Ze^2}{\xi^2 + \eta^2} + eF \cdot \frac{\xi^2 - \eta^2}{2}.$$

ϕ occurs "cyclically" in (3). This suggests to us to assume

$$\psi = \Phi(\xi, \eta) e^{\pm im\phi}, \quad m \geq 0.$$

From (3) we then get, after division by $\xi \eta$,

$$\left. \begin{aligned} & \frac{1}{\xi} \frac{\partial}{\partial \xi} \left(\xi \frac{\partial \Phi}{\partial \xi} \right) - \frac{m^2}{\xi^2} \Phi + \frac{8\pi^2\mu}{h^2} \left\{ E\xi^2 + Ze^2 - \frac{eF}{2} \xi^4 \right\} \Phi \\ & + \frac{1}{\eta} \frac{\partial}{\partial \eta} \left(\eta \frac{\partial \Phi}{\partial \eta} \right) - \frac{m^2}{\eta^2} \Phi + \frac{8\pi^2\mu}{h^2} \left\{ E\eta^2 + Ze^2 + \frac{eF}{2} \eta^4 \right\} \Phi = 0 \end{aligned} \right\} \quad (4)$$

This method of writing the equation allows us to recognise the possibility of carrying the separation still further:

$$\Phi = f_1(\xi) f_2(\eta).$$

* Schrödinger uses the co-ordinates $\lambda_1 = \xi^2, \lambda_2 = \eta^2, \phi = \phi$.

If β is a separation constant, (4) sub-divides into

$$\frac{1}{\xi} \frac{d}{d\xi} \left(\xi \frac{df_1}{d\xi} \right) + \frac{8\pi^2\mu}{h^2} \left\{ E\xi^2 + Ze^2 - \beta - \frac{m^2 h^2}{8\pi^2 \mu \xi^2} - \frac{eF}{2} \xi^4 \right\} f_1 = 0. \quad (5a)$$

and

$$\frac{1}{\eta} \frac{d}{d\eta} \left(\eta \frac{df_2}{d\eta} \right) + \frac{8\pi^2\mu}{h^2} \left\{ E\eta^2 + Ze^2 + \beta - \frac{m^2 h^2}{8\pi^2 \mu \eta^2} + \frac{eF}{2} \eta^4 \right\} f_2 = 0. \quad (5b)$$

Eqns. (5) may be treated together further. If, linking up with eqns. (19), (20a), (20b) of I, Note 11, we introduce the abbreviations

$$\left. \begin{aligned} r &= \begin{cases} \xi^2, & f = f_1, \\ \eta^2, & f = f_2, \end{cases} \quad A = \frac{2\pi^2\mu}{h^2} E \\ B &= \frac{\pi^2\mu}{h^2} (Ze^2 \mp \beta), \quad C = -\frac{m^2}{4}, \quad \lambda = \pm \frac{\pi^2\mu}{h^2} eF \end{aligned} \right\} \quad (6)$$

and combine (5a, b), we get,

$$\frac{d^2 f}{dr^2} + \frac{1}{r} \frac{df}{dr} + \left(A + \frac{2B}{r} + \frac{C}{r^2} \right) f = \lambda r f \quad (7)$$

The *unperturbed* equation ($\lambda = 0$, Kepler problem in parabolic co-ordinates) is analogous to eqn. (3), § 7, Chap. I. Its integration need therefore only be indicated and we shall restrict ourselves to the case $E < 0$, $A < 0$, that is, to the discrete spectrum:

asymptotic solution: $f = e^{-\sqrt{-A}r}$,

unnamed variable: $\rho = 2\sqrt{-A}r$,

assumption for f : $f = e^{-\frac{\rho}{2}v}$,

differential equation for v :

$$v'' + \left(\frac{1}{\rho} - 1 \right) v' + \left\{ \left(-\frac{1}{2} + \frac{B}{\sqrt{-A}} \right) \frac{1}{\rho} + \frac{C}{\rho^2} \right\} v = 0 \quad (8)$$

expansion for the neighbourhood of $\rho = 0$: $v = \rho^\gamma \Sigma a_\nu \rho^\nu$,

characteristic equation for γ : $\gamma^2 = -C$, and hence by (6),

$$\gamma = \pm \frac{m}{2},$$

recurrence formula for the α 's:

$$\alpha_{\nu+1} \{ \quad \} + \alpha_\nu \left\{ -\left(\frac{m}{2} + \nu \right) - \frac{1}{2} + \frac{B}{\sqrt{-A}} \right\} = 0,$$

polynomial condition, breaking off the recurrence for $\nu = n_i$:

$$\frac{B}{\sqrt{-A}} = \frac{1}{2} + \frac{m}{2} + n_i \quad (9)$$

n_i represents the two values n_1 and n_2 , which respectively belong to f_1 and f_2 , just as B represents two values. By adding the two equations (9) for $i = 1, 2$ we get (remembering that by (6) A is one-valued),

$$\frac{2\pi^2\mu}{h^2} \frac{Ze^2}{\sqrt{-A}} = 1 + m + n_1 + n_2 = n \quad (10)$$

156 Chapter II. Perturbations. Diffraction Problems

We call the right-hand side the "principal quantum number" n . By squaring and substituting for A we get the Balmer formula:

$$E = -\frac{2\pi^2\mu e^4 Z^2}{h^2 n^2} = -\frac{R h Z^2}{n^2} \quad . \quad . \quad . \quad (11)$$

So far we have only shown that the unperturbed Kepler problem allows itself to be integrated just as well in parabolic co-ordinates as in polar co-ordinates and leads to the same result, namely the Balmer formula. All this is in agreement with the earlier procedure of I, Chap. V, § 5, p. 285. According to the formulæ just above the components f_1, f_2 of the unperturbed proper functions are given by

$$f = e^{-\frac{\rho}{2}} \rho^{\frac{m}{2}} w, \quad \text{where } w = \sum_0^{n_i} \alpha_i \rho^i \quad . \quad . \quad . \quad (12)$$

The differential equation for w follows from that for v , eqn. (8), if we substitute the values of C from (6) and of $B/\sqrt{-A}$ from (9):

$$\rho w'' + (m+1-\rho)w' + n_i w = 0.$$

Comparing this with (12a) of § 7, Chap. I (differential equation of the derivatives of Laguerre's polynomial), we see that

$$w = \frac{d^m}{d\rho^m} L_m + n_i(\rho) \quad . \quad . \quad . \quad (13)$$

B. Perturbation of the Proper Values to the First Order

We now turn to the perturbation equation (7) which we simplify by introducing $\rho = 2\sqrt{-A}r$, and get

$$f'' + \frac{1}{\rho}f' + \left(-\frac{1}{4} + \frac{B}{\sqrt{-A}}\frac{1}{\rho} - \frac{m^2}{4\rho^2}\right)f = \lambda'\rho f \quad . \quad . \quad (14)$$

Here we have set [cf. eqn. (6)]

$$\lambda' = \frac{\lambda}{(2\sqrt{-A})^2}, \quad C = -\frac{m^2}{4} \quad . \quad . \quad . \quad (15)$$

By (6) the sign of λ' is positive or negative according as f denotes the separation function f_1 or f_2 . In the same way, by (6), $B/\sqrt{-A}$ has a different meaning according to whether f_1 or f_2 is being discussed. This quantity $B/\sqrt{-A}$ represents the *proper value* of our eqn. (14) and is to be denoted by the abbreviation η .

To be able to apply the general method of the preceding section directly, we must first make our perturbation equation (14) self-adjoint; this is easily done by multiplying it by ρ . This transforms (14) into

$$\frac{d}{d\rho}(\rho f') + \left(\eta - \frac{\rho}{4} - \frac{m^2}{4\rho}\right)f = \lambda'\rho^2 f \quad . \quad . \quad . \quad (14a)$$

The first approximation of the proper value η was determined in (9) and is to be called η_0 ; on account of the occurrence of the parabolic quantum number n_i in (9) η_0 is also different for f_1 and f_2 . We next expand η and f in the sense of the previous section in terms of the perturbation parameter (15)

$$\eta = \eta_0 + \lambda' \epsilon, \quad f = f_0 + \lambda' \phi \quad (16)$$

where f_0 is given by (12)

Substituting (16) in (14a) and neglecting the term in λ'^2 , we get for ϕ the equation

$$\frac{d}{d\rho}(\rho\phi') + \left(\eta - \frac{\rho}{4} - \frac{m^2}{4\rho}\right)\phi = (\rho^2 - \epsilon)f_0 \quad (16a)$$

Hence it follows from eqn. (8) of § 1 (in our case $s = \rho^2$, $p = 1$, $dq = d\rho$)

$$\epsilon = \frac{K}{J}, \quad K = \int_0^\infty \rho^2 f_0^2 d\rho, \quad J = \int_0^\infty f_0^2 d\rho \quad (17)$$

In calculating J and K we proceed on the lines of eqn. (17), § 7, Chap. I. In consequence of our present eqns. (12) and (13) we have

$$f_0^2 = e^{-\rho} \rho^{mw} \cdot \frac{d^m L_\nu}{d\rho^m},$$

where we have used the abbreviation $\nu = m + n_i$. Thus both integrals J and K may be written in the form

$$\int_0^\infty e^{-\rho} G \frac{d^m L_\nu}{d\rho^m} d\rho \quad (18)$$

where G is an integral function of degree ν in the case of J , and of degree $\nu + 2$ in the case of K :

$$J) \quad G = \rho^{mw} = a\rho^\nu + \dots$$

$$K) \quad G = \rho^{m+2w} = a\rho^{\nu+2} + b\rho^{\nu+1} + c\rho^\nu + \dots$$

To calculate J we perform integration by parts m times on (18); we shall see that only $e^{-\rho}$ need be differentiated:

$$J = \int_0^\infty e^{-\rho} (a\rho^\nu + \dots) L_\nu d\rho.$$

If we now substitute for L_ν the expression (14) from § 7, Chap. I, the exponential function cancels out and we get

$$J = \int_0^\infty (a\rho^\nu + \dots) \frac{d^\nu}{d\rho^\nu} (\rho^\nu e^{-\rho}) d\rho.$$

We may integrate this once again by parts, which makes all terms with exponents $< \nu$ drop out:

$$J = (-1)^\nu \nu! a \int_0^\infty \rho^\nu e^{-\rho} d\rho = (-1)^\nu (\nu!)^2 a.$$

In a fully analogous manner we find that

$$K = (-1)^\nu (\nu!)^2 [c + (\nu+1)(\nu+1-m)b + \frac{1}{2}(\nu+1)(\nu+2)(\nu+1-m)(\nu+2-m)a].$$

Hence

$$\left. \begin{aligned} \frac{K}{J} &= \frac{c}{a} + (\nu + 1)(\nu + 1 - m)\frac{b}{a} \\ &+ \frac{1}{2}(\nu + 1)(\nu + 2)(\nu + 1 - m)(\nu + 2 - m) \end{aligned} \right\} \quad (19)$$

To determine a , b , and c here we must revert to eqn. (29) of § 7, Chap. I; we get by differentiating m times

$$\begin{aligned} w = \frac{d^m L_\nu}{d\rho^m} &= (-1)^m \left(\frac{\nu!}{(\nu - m)!} \rho^{\nu - m} - \frac{\nu^2(\nu - 1)!}{(\nu - m - 1)!} \rho^{\nu - m - 1} \right. \\ &\quad \left. + \frac{1}{2} \frac{\nu^2(\nu - 1)^2(\nu - 2)!}{(\nu - m - 2)!} \rho^{\nu - m - 2} + \dots \right). \end{aligned}$$

From this we obtain the ratios of our coefficients a , b , c in G as

$$\frac{c}{a} = \frac{1}{2}\nu(\nu - 1)(\nu - m)(\nu - m - 1), \quad \frac{b}{a} = -\nu(\nu - m).$$

Substituting these values in (19) gives

$$\frac{K}{J} = 6\nu^2 + 6\nu(1 - m) + m^2 - 3m + 2 \quad (20)$$

By (17) this is the perturbation of the proper value ϵ ; and we must distinguish between ϵ_1 and ϵ_2 according as we are dealing with f_1 or f_2 . In the sequel we require only the difference $\epsilon_1 - \epsilon_2$, which, by (20), comes out as

$$\epsilon_1 - \epsilon_2 = 6(\nu_1 - \nu_2)(\nu_1 + \nu_2 + 1 - m).$$

On account of the meaning of $\nu_i = m + n_i$ and of $n = 1 + m + n_1 + n_2$ this equation is identical with

$$\epsilon_1 - \epsilon_2 = 6(n_1 - n_2)n \quad (21)$$

C. The Stark Effect of the First Order

We now write down the perturbed proper value η for both parabolic co-ordinates according to (16), substituting for η_0 from (9) and for λ' from (15) [concerning the sign of $|\lambda|$, cf. (6)]:

$$\left. \begin{aligned} \eta_1 &= \frac{1}{2} + \frac{m}{2} + n_1 + \frac{|\lambda|}{(2\sqrt{-A})^3} \epsilon_1 \\ \eta_2 &= \frac{1}{2} + \frac{m}{2} + n_2 - \frac{|\lambda|}{(2\sqrt{-A})^3} \epsilon_2 \end{aligned} \right\} \quad (22)$$

Now $\eta = B/\sqrt{-A}$ and hence by (16)

$$\left. \begin{aligned} \eta_1 &= \frac{\pi^2 \mu}{h^2 \sqrt{-A}} (Ze^2 - \beta) \\ \eta_2 &= \frac{\pi^2 \mu}{h^2 \sqrt{-A}} (Ze^2 + \beta) \end{aligned} \right\} \quad (23)$$

Therefore, by adding the two eqns. (22) we get

$$\frac{2\pi^2\mu}{h^2} \frac{Ze^2}{\sqrt{-A}} = n + \frac{|\lambda|}{(2\sqrt{-A})^3}(\epsilon_1 - \epsilon_2),$$

or, on account of (21),

$$\frac{2\pi^2\mu}{h^2} \frac{Ze^2}{\sqrt{-A}} = n \left(1 + \frac{6|\lambda|(n_1 - n_2)}{(2\sqrt{-A})^3} \right).$$

Neglecting λ^2 , which is reasonable, we get from this

$$- \left(\frac{h^2}{2\pi^2\mu} \right)^2 \frac{A}{(Ze^2)^2} = \frac{1}{n^2} \left(1 - \frac{3|\lambda|(n_1 - n_2)}{(\sqrt{-A})^3} \right).$$

On the left-hand side we substitute A from (6); on the right we insert the 0th approximation of A from (10) and obtain

$$- \frac{h^2 E}{2\pi^2\mu Z^2 e^4} = \frac{1}{n^2} - \frac{3h^6 |\lambda| (n_1 - n_2) n}{16\pi^6 \mu^3 Z^3 e^6}.$$

If finally we take into account the significance of λ in eqn. (6) and of the Rydberg constant R , we have

$$E = - \frac{R h Z^2}{n^2} + \frac{3 h^2 F}{8 \pi^2 \mu e Z} (n_1 - n_2) n. \quad (24)$$

This is exactly the Schwarzschild-Epstein formula for the Stark effect of the first order, which we discussed in I, Chap. V, § 5, and compared with the results of experiment; see in particular eqn. (1) of that section.

Not only is the result the same, however, but also the method by which we have arrived at it runs parallel with the earlier method in I, Note 11, p. 587. For example, here, as there, we had to eliminate the separation constant β by adding the two values of $B/\sqrt{-A}$. The calculation of A from the sum so obtained also ran along the same lines. The only difference is that earlier we determined $B/\sqrt{-A}$ by obtaining an approximation to a complex integral, whereas here we determined it from a differential equation which was integrated approximately by means of the theory of perturbations.

We have here occupied ourselves explicitly only with the energy E , which we compounded from the perturbed proper values η of our two differential equations, and we have eliminated the separation-constant β . But our calculation also contains the perturbation-correction of β implicitly, which may be taken directly, according to eqns. (23), from the perturbed values of η .

In order to connect this method with the general rules of the preceding section, we make the following remarks. The Kepler problem is degenerate: therefore the perturbed Kepler problem is presumably to be treated according to § 1 B. Instead of this we have preferred to separate the variables in the perturbed problem and to apply the simpler

methods of § 1 A. This possibility was based on a general theorem drawn from the theory of the Sturm-Liouville boundary problems: the proper values in the case of ordinary differential equations of this type are always simple. It is in accordance with this that we started out in our calculation of perturbations not from the whole energy E involved in the problem, but from the proper values η of the separated differential equations and corrected these values by adding perturbation terms. Only at the end, in eqn. (24), did we pass on to the total energy E of the problem.

D. Stark Effect of Higher Order

We make the assumption (16) complete by developing the proper value and the proper function in a series of powers of the perturbation parameter:

$$\eta = \sum \lambda'^k \eta_k, \quad f = \sum \lambda'^k f_k. \quad (25)$$

We substitute (25) in (14a). By equating the various powers of λ' to zero we get a system of equations, giving recurrence formulæ for the f_k 's. The first of them is eqn. (16a) in somewhat different notation; the "0th" is the differential equation of the unperturbed Kepler problem in parabolic co-ordinates. The k th equation of the system is

$$\frac{d}{d\rho}(\rho f'_k) + \left(\eta_0 - \frac{\rho}{4} - \frac{m^2}{4\rho}\right)f_k = (\rho^2 - \eta_1)f_{k-1} - \eta_2 f_{k-2} \dots - \eta_k f_0.$$

The case $k = 2$ interests us:

$$\frac{d}{d\rho}(\rho f'_2) + \left(\eta_0 - \frac{\rho}{4} - \frac{m^2}{4\rho}\right)f_2 = (\rho^2 - \eta_1)f_1 - \eta_2 f_0,$$

The solution of the corresponding homogeneous equation (left side of the preceding equation set equal to zero) is f_0 . The right side must be "orthogonal" to this solution [analogously to eqn. (7a) of § 1]. In our present notation this means that

$$\int \{(\rho^2 - \eta_1)f_1 - \eta_2 f_0\}f_0 d\rho = 0,$$

that is,

$$J\eta_2 = \int \rho^2 f_1 f_0 d\rho - \eta_1 \int f_0 f_1 d\rho \quad (26)$$

J is the integral taken from eqn. (17). To evaluate the right-hand side we must first know f_1 . f_1 is identical with the function ϕ in eqn. (16a) and is calculated according to eqn. (11) of § 1. Hence, following rule (9) of § 1, we expand the right-hand side of (16a) in terms of all the proper functions f_0 of the unperturbed equation; in doing so we take into account that in our case the weight function $p = 1$, and that we have now to write η_1 in place of ϵ :

$$(\rho^2 - \eta_1)f_0 = \sum A_i f_{0i} \quad (27)$$

On the left-hand side f_0 , or more precisely f_{0n} , denotes the proper function of zero approximation, from which we started, that is the proper function with the parabolic quantum number n_1 or n_2 . On the right-hand side f_{0i} denotes all the remaining proper functions of zero approximation with $i \geq n$. By (11) of § 1 we therefore represent f_1 thus:

$$f_1 = \sum_{(i)}' \frac{A_i f_{0i}}{n - i} \quad (28)$$

[the difference of proper values that occurs in eqn. (11), § 1, here simply becomes $n - i$ by eqn. (9)]. From this we see that the second integral on the right-hand side of (26) vanishes. For here f_0 denotes as much as f_{0n} ; since in (28) the term with $i = n$ drops out and all integrals $\int f_{0i} f_{0n} d\rho$ vanish for $i \neq n$ on account of orthogonality, we also get

$$\int f_0 f_1 d\rho = 0.$$

Substituting (28) in (26)

$$J\eta_2 = \sum_{(i)}' \frac{A_i}{n - i} \int \rho^2 f_{0n} f_{0i} d\rho \quad (29)$$

From (27) it follows, however, in the manner of Fourier, that

$$\int \rho^2 f_{0n} f_{0i} d\rho - \eta_1 \int f_{0n} f_{0i} d\rho = A_i \int f_{0i}^2 d\rho \quad (29a)$$

Again, the second member on the left vanishes on account of the condition of orthogonality; we shall call the factor of A_i on the right J_i , so distinguishing it from the previous $J = \int f_{0n}^2 d\rho$, for which we shall from now on write J_n for clearness. If we denote the first integral by Y_i , eqn. (29a) then asserts

$$Y_i = \int \rho^2 f_{0n} f_{0i} d\rho = A_i J_i \quad (29b)$$

If we substitute this in (29), we get

$$\eta_2 = \sum_{(i)}' \frac{A_i^2}{n - i} \cdot \frac{J_i}{J_n} \quad (30)$$

It is now easily seen that $A_i = 0$, if

$$i < n - 2 \text{ or } i > n + 2$$

and so only the four following coefficients of A come into consideration:

$$A_{n-2}, A_{n-1}, A_{n+1}, A_{n+2}.$$

For if we write the integral (29b) in the form (18) and set $\nu = m + n$, the integral function $G = \rho^{m+2} w_i$ becomes of degree $m + i + 2$ and this degree becomes $< \nu$, if $i < n - 2$. In performing the integrations by parts as before, all the terms of G vanish. The same is true for $i > n + 2$,

if we allow i and n to change rôles. By this method we also calculate the values of Y_i for $i = n \mp 2$ and $i = n \mp 1$, and hence also, by (29b), $A_i = Y_i/J_i$.

We shall at once write down the successive terms of the sum (30):

$$\begin{aligned} \frac{A_i^2}{n-i} \cdot \frac{J_i}{J_n} &= \frac{1}{2} \nu(\nu-1)(\nu-m)(\nu-1-m) & i = n-2, \\ &= 4\nu(\nu-m)(2\nu-m)^2 & i = n-1, \\ &= -4(\nu+1)(\nu+1-m)(2\nu+2-m)^2 & i = n+1, \\ &= -\frac{1}{2}(\nu+2)(\nu+1)(\nu+2-m)(\nu+1-m) & i = n+2. \end{aligned}$$

From these we get by addition the value η_2 of the perturbation for the individual parabolic co-ordinate. n is to be replaced by n_1 or n_2 if we deal with ξ or η , respectively. The sum of the η_2 's for the two co-ordinates becomes (the principal quantum number $n = n_1 + n_2 + m + 1$ can be taken outside the bracket)

$$\left. \begin{aligned} \sum \eta_2 &= -2m\{4m^2 + 17(n_1 + n_2 + 1)m \\ &\quad + 34(n_1^2 + n_2^2 - n_1n_2) + 17(n_1 + n_2) + 18\} \end{aligned} \right\} \quad (31)$$

In (25) η denoted the proper value $B/\sqrt{-A}$ for the individual parabolic co-ordinate. If we again denote the sum of both by $\sum \eta$, then by (6)

$$\sum \eta = \frac{2\pi^2\mu}{h^2} \frac{Ze^2}{\sqrt{-A}} \quad (32)$$

On the other hand, the expansion (25) gives us, if we express λ' in terms of λ by (15) and substitute n for $\sum \eta_0$ by (9),

$$\sum \eta = n + \frac{\lambda}{(2\sqrt{-A})^3}(\epsilon_1 - \epsilon_2) + \frac{\lambda^2}{(2\sqrt{-A})^6} \sum \eta_2 \quad (33)$$

In this expression $\epsilon_1 - \epsilon_2$ is given by (21). In the second term on the right we put for $\sqrt{-A}$ the value of its first approximation from (23a), and for the third member on the right the approximation of 0th degree from (10). By comparing (32) and (33) we then get the value of A to a *second* approximation, and hence by (6) also that of the energy E . The Stark effect of the second order is represented by the term with F^2 in the expression for E . We get for this

$$-\frac{\hbar^6 F^2}{16(2\pi e)^6 \mu^3 Z^4} n^4 [17n^2 - 3(n_1 - n_2)^2 - 9m^2 + 19] \quad (34)$$

This value was obtained simultaneously * by Wentzel and Waller. Neither used the perturbation theory of Schrödinger. Wentzel makes his calculation by the simple and interesting method which we have given in § 12, Chap. I. Waller exploits the "polynomial method" to the fullest extent

* G. Wentzel, *Zeitschr. f. Phys.*, **38**, 518 (1927); J. Waller, *ibid.*, p. 635.

even for the solutions of higher order. [As we saw earlier, the 0th approximation f_0 is, except for the factor $e^{-\rho/2}\rho^{m/2}$ of eqn. (12), a polynomial of degree n_i , the 1st approximation f_1 is a polynomial of degree $n_i + 2$, and, correspondingly, the 2nd approximation is a polynomial of degree $n_i + 4$, and so forth.]

As indicated in I, Note 11, p. 594, the Stark effect of the second order has already been calculated by Epstein on the *old* quantum theory. In appearance his formula differs from (34) only in the absence of the constant member (19) in the brackets; then there is the more fundamental difference that the quantum number of Epstein corresponding to our m assumes the values 1, 2, . . . and excludes zero, where with us m assumes the values 0, 1, 2, . . . The author detected in observations made by Takamine and Kokubu a slight deviation from Epstein's formula. Recently very accurate experiments by H. Rausch von Traubenberg and R. Gebauer confirm the formula (34) perfectly, being at the same time in contradiction with Epstein's earlier formula.*

The difference between the wave-mechanical and Epstein's formula is relatively greater for small quantum numbers than for great quantum numbers. The observations just mentioned come among the high quantum numbers H_γ , H_β (it is the quantum number of the initial state that counts). In the ground state ($n = 1$, $n_1 = n_2 = m = 0$, or, in Epstein's case, $n_1 = n_2 = 0$, $m = 1$) the difference between (34) and Epstein's value is given by the ratio $36 : 8 = 4.5 : 1$. We consider two examples: 1. A hydrogen or alkali atom in the ground state under the influence of an alternating electric field; the Stark effect of the first order drops out when averaged over the time, and the Stark effect of the second order gives the "polarisation" or "deformation" of the atom. 2. Core of the helium atom in the field of the excited outer electron; on account of the changing direction of the field, here too the Stark effect of the second order alone comes into question and gives the polarisation of the He core. In both cases the effect calculated by wave-mechanics is 4.5 times as great as that expected according to the old theory. For details see the papers by Wentzel and Waller, *loc. cit.*

Of course the new theory is superior to the old theory chiefly in allowing us to calculate, besides the wave-lengths of the components of the Stark effect, also their intensities. Schrödinger has done this to a sufficient degree of approximation (proper functions of 0th approximation, wave-lengths to the first approximation), and has met with success in that his values for the intensities for most components come much nearer to the experimental values than those estimated by means of the correspondence principle.† A more exact calculation of the intensity (proper functions

* Naturwissenschaften, August, 1928, and June, 1929; *Zs. f. Phys.*, 56, 254 (1929).

† According to Mark and Wierl we can choose the conditions (field perpendicular to the direction of motion), so that the intensities calculated by Schrödinger at any rate come out to the right order qualitatively. Cf. Naturwissenschaften, 1928.

to a first approximation) produces no appreciable change for the fields that are experimentally possible, in fact only an unobservably small asymmetry in the intensity scheme for the right and left sides.*

§ 3. Theory of Dispersion

In the older quantum theory we had no means of treating the forced radiation of atoms; even free radiation could be dealt with only with the help of a special axiom (cf. the beginning of I, Chap. I). That is why we did not discuss the theory of dispersion in the earlier editions of *Atombau*. In contradistinction to this, questions of dispersion, just like those of free emission, fit quite naturally and without constraint into the system of wave-mechanics.

Schrödinger in his "fourth communication" (on Quantisation as a Problem of Proper Values) treats the theory of dispersion as a pulsating Stark effect by introducing an electric potential $V(t)$ varying in time into the wave-equation. Instead of this we may follow O. Klein † and describe the light wave by means of a vector potential. Both methods bring us to the same goal provided we make the assumption (which is, of course, fully justified for optical purposes) that the potential gradient may be neglected inside the limits of the atom. We shall adopt the second method here as it enables us simultaneously to prepare for the treatment of the photo-electric effect in the next section.

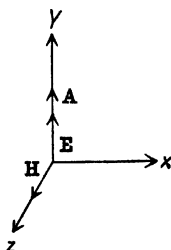


FIG. 16.

A. Perturbation of the Atom by an Incident Light-Wave

Suppose the incident light-wave to come from the negative x -direction and the electric force to be polarised in the y -direction. The following assumption for the electrodynamic potentials \mathbf{A} and ϕ corresponds with these conditions (see Fig. 16):

$$\mathbf{A} = \mathbf{A}_y = a \cos \omega \left(t - \frac{x}{c} \right), \quad \phi = 0 \quad . \quad . \quad (1)$$

By the familiar rules

$$\mathbf{H} = \text{curl } \mathbf{A}, \quad \mathbf{E} = -\text{grad } \phi - \frac{1}{c} \frac{\mathbf{A}}{t},$$

it follows from (1) that

$$\left. \begin{aligned} \mathbf{H}_x = \mathbf{H}_y = 0, \quad \mathbf{H}_z &= \frac{a\omega}{c} \sin \omega \left(t - \frac{x}{c} \right) \\ \mathbf{E}_x = \mathbf{E}_z = 0, \quad \mathbf{E}_y &= \frac{a\omega}{c} \sin \omega \left(t - \frac{x}{c} \right) \end{aligned} \right\} \quad . \quad . \quad (1a)$$

and so, in fact, represents a field of waves (*Wellenfeld*) of the desired kind.

* According to unpublished calculations by W. Zimmermann, Munich, 1928.

† Zeitschr. f. Phys., 41, 407 (1927).

We insert the values of \mathbf{A} from (1) in the wave-equation (10) of § 9' Chap. I, which contains the time :

$$\left. \begin{aligned} \Delta u - \frac{4\pi im}{h} \frac{\partial u}{\partial t} - \frac{8\pi^2 m}{h^2} (E_0 + U)u \\ = - \frac{4\pi ie}{hc} a \cos \omega \left(t - \frac{x}{c} \right) \frac{\partial u}{\partial y} \end{aligned} \right\} \quad (2)$$

This equation holds for the single electron in the hydrogen atom (m = the mass of the electron) to which we can restrict our attention here, although our final result claims a far more general significance. It is convenient to split up the cosine into its exponential components; we then write (2) in the form

$$\left. \begin{aligned} \Delta u - \frac{4\pi im}{h} \frac{\partial u}{\partial t} - \frac{8\pi^2 m}{h^2} (E_0 + U)u \\ = \rho \frac{\partial u}{\partial y} \left\{ e^{i\omega \left(t - \frac{x}{c} \right)} + e^{-i\omega \left(t - \frac{x}{c} \right)} \right\} \end{aligned} \right\} \quad (3)$$

ρ is the perturbation parameter

$$\rho = - \frac{2\pi ie}{hc} a; \quad (4)$$

since the amplitude a of the vector-potential \mathbf{A} occurs in ρ , ρ can be treated as an arbitrarily small quantity.

Let the unperturbed state of the atom and the state perturbed by the light-wave be respectively represented by

$$u_k = \psi_k e^{\frac{2\pi i}{h} E_k t}, \quad u = u_k + \rho w \quad (5)$$

u_k satisfies eqn. (3) with $\rho = 0$; by substituting for u in (3) we get for w the following differential equation, if we neglect a term in ρ^2 :

$$\left. \begin{aligned} \Delta w - \frac{4\pi im}{h} \frac{\partial w}{\partial t} - \frac{8\pi^2 m}{h^2} (E_0 + U)w \\ = \frac{\partial \psi_k}{\partial y} \left\{ e^{\frac{2\pi i}{h} E_k t + i\omega \left(t - \frac{x}{c} \right)} + e^{\frac{2\pi i}{h} E_k t - i\omega \left(t - \frac{x}{c} \right)} \right\} \end{aligned} \right\} \quad (6)$$

As the right-hand side depends on t we assume the following form for w :

$$w = w_+ \cdot e^{\frac{2\pi i}{h} (E_k + h\nu)t} + w_- \cdot e^{\frac{2\pi i}{h} (E_k - h\nu)t} \quad (7)$$

ν denotes vibration number corresponding to the circular frequency ω :
 $\nu = \omega/2\pi$. If we also introduce the associated wave-length

$$\lambda = c/\nu = 2\pi c/\omega$$

(6) gives us the following equation for determining w_{\pm} :

$$\Delta w_{\pm} + \frac{8\pi^2 m}{h^2} (E_k \pm h\nu - E_0 - U)w_{\pm} = \frac{\partial \psi_k}{\partial y} \cdot e^{\mp \frac{2\pi i}{\lambda} x} \quad (8)$$

Whereas in the theory of perturbations of § 1 and its application to the Stark effect not only the proper *functions* but also the proper *values* were changed by the perturbation [cf., for example, eqn. (4a) of § 1], we had only to correct the proper *functions* of the atom in eqn. (5) of the present section. The reason for this is that here we started out from the time-equation (3) in which no proper value parameter occurs at all. The proper value E_k does not occur as a parameter in the perturbation problem in this method of treatment, but only as a parameter of the unperturbed atomic state. In spite of this difference § 1 indicates the lines along which we must proceed for treating equation (8) further: namely, we must expand the right-hand side of (8), as in (9) of § 1, in terms of the proper functions ψ_j of the unperturbed problem:

$$\frac{\partial \psi_k}{\partial y} \cdot e^{\mp 2\pi i \frac{x}{\lambda}} = \sum A_j \psi_j \quad . \quad . \quad . \quad (9)$$

These proper functions ψ_j satisfy the well-known equation

$$\Delta \psi_j + \frac{8\pi^2 m}{h^2} (E_j - E_0 - U) \psi_j = 0 \quad . \quad . \quad . \quad (8a)$$

when we take into account our present notation for the potential energy ($V = E_0 + U$). If we assume the solution in the form

$$w_{\pm} = \sum B_j \psi_j$$

it follows from (8) and (8a), in the manner of eqn. (10a) in § 1 (the constant $\frac{8\pi^2 m}{h^2}$ corresponds in a certain sense to the "weight-function" p used earlier),

$$B_j = \frac{h^2}{8\pi^2 m} \cdot \frac{A_j}{E_k - E_j \pm h\nu}$$

and hence

$$w_{\pm} = \frac{h^2}{8\pi^2 m} \sum^{(j)} \frac{A_j \psi_j}{E_k - E_j \pm h\nu} \quad . \quad . \quad . \quad (10)$$

Since, in contradistinction to § 1, there is here no reason why A_j should vanish in the case $j = k$, the summation in (10) must be performed over all values of j , smaller, equal to or greater than k . We repeat again that the index k denotes the original state of the atom, in which it is struck by the light wave. By eqns. (4), (5), (7), and (10) the perturbed state is represented by

$$u = \psi_k e^{\frac{2\pi i}{h} E_k t} - \frac{hi}{4\pi c} \frac{e}{m} a \left\{ \sum^{(j)} \frac{A_j^+ \psi_j}{E_k - E_j + h\nu} e^{\frac{2\pi i}{h} (E_k + h\nu)t} + \sum^{(j)} \frac{A_j^- \psi_j}{E_k - E_j - h\nu} e^{\frac{2\pi i}{h} (E_k - h\nu)t} \right\} \quad (11)$$

The symbol A_j^{\pm} indicates that the difference which, by (9) exists, strictly speaking, between the values of A_j in the two sums has been taken into

account; but if we neglect the size of the atom in comparison with the wave-length λ , we may regard the two A_j 's as being the same; we shall do this in the next section.

Eqn. (11) shows that under the influence of the light-wave not only the original state k , but also all other states j for which A_j does not just happen to be equal to zero are excited. If besides the discrete spectrum there is also a continuous spectrum of proper values, the summation terms in (11) of course become supplemented by corresponding integrals, cf. the end of § 1 A and § 4.

B. Frequency and Intensity of the Resulting Vibrations The Dispersion Formula

In this section we shall follow § 5 of Chap. I closely. For by the eqn. (16) that occurs there we calculate the *electric density* of our distribution of state u (*Zustandsverteilung*), that is the quantity uu^* . From (11) we next get, using an abbreviated form,

$$uu^* = \left(\psi_k e^{\frac{2\pi i}{h} E_k t} - \frac{hi}{4\pi c} \frac{e}{m} a \left\{ \begin{array}{c} \\ \end{array} \right\} \right) \\ \left(\psi_k^* e^{-\frac{2\pi i}{h} E_k t} + \frac{hi}{4\pi c} \frac{e}{m} a \left\{ \begin{array}{c} \\ \end{array} \right\}^* \right).$$

In performing the multiplication we can obviously neglect the term in a^2 , as it is a quantity of the second order. We get

$$uu^* = \psi_k \psi_k^* + \frac{hi}{4\pi c} \frac{e}{m} a \left\{ e^{+i\omega t} \sum^+ + e^{-i\omega t} \sum^- \right\} \quad (12)$$

with the abbreviations

$$\sum^\pm = \sum_{(j)} \left(\frac{A_j^* \psi_j^* \psi_k}{E_k - E_j \mp h\nu} - \frac{A_j \psi_j \psi_k^*}{E_k - E_j \pm h\nu} \right). \quad (12a)$$

From (12) we see that *the density performs forced vibrations with the frequency ω of the incident light-wave; the original frequency of the atom, given by E_k , has vanished from (12)*. In this way we have deduced a fundamental characteristic of all dispersion phenomena (cf. the remarks in D relating to the Smekal transitions).

If, in particular, ψ is real or if, as in the case of the Kepler problem in the absence of a magnetic field, it can be written in the real form, and if, as agreed, we neglect the exponential factor in (9), so that A_j also becomes real, then we can verify by (12a) that $\sum^- = -\sum^+$. From this it follows that that component of the density which contains the time varies as $\sin \omega t$, that is, vibrates *in phase with* the electric intensity of field E_y , eqn. (1a), if we calculate the latter for the middle point $x = 0$, of the atom.

This, too, is in conformity with the fundamental laws of the theories of dispersion.

The simplified value of $\sum +$ that results from these assumptions is

$$\sum + = 2h\nu \sum \frac{A_j \psi_j \psi_k}{(E_k - E_j)^2 - (h\nu)^2} \quad . \quad . \quad . \quad (13)$$

and eqn. (12) gives us, when we pass on to the electric density $\rho = euv^*$,

$$\rho = e\psi_k^2 - \frac{h^2\nu}{\pi c} \frac{e^2}{m} \sum_j \frac{A_j \psi_j \psi_k}{(E_k - E_j)^2 - (h\nu)^2} \sin \omega t \quad . \quad . \quad (14)$$

We next enquire into the *intensity* of the emitted radiation. To find this we must, by § 5, Chap. I, form the *electric moment* of our density of distribution. Our first concern is with the moment for any arbitrary co-ordinate direction q , namely (cf. (17) in § 5, Chap. I) :

$$M = \int q\rho d\tau \quad . \quad . \quad . \quad . \quad (15)$$

(We may, in particular, choose the direction q parallel to the direction of the exciting electric force, thus, as the y -direction, cf. Fig. 16, as we shall do presently in linking up with the classical theory of dispersion.) We omit the part ψ_k^2 which is independent of the time as being of no account, and we effect the integration prescribed in (15) term by term for the numerator of the sum (14). We write

$$\int q\psi_j\psi_k d\tau = q_{jk} \quad . \quad . \quad . \quad . \quad (16)$$

Here, q_{jk} has the same meaning as in eqn. (21), § 5, Chap. I, and is thus a measure of the probability of transition from the state k to the state j or *vice versa*. Hence we get from (14), (15) and (16) as the variable component of M if we complete the suffix of A_j by writing it as A_{jk} :

$$M = - \frac{h^2\nu}{\pi c} \frac{e^2}{m} \sum_j \omega \frac{A_{jk}q_{jk}}{(E_k - E_j)^2 - (h\nu)^2} \sin \omega t \quad . \quad . \quad (17)$$

We write $E_j - E_k = h\nu_{jk}$, so that ν_{jk} denotes the vibration number of the *free* radiation emitted in the transition $j \rightarrow k$, in contradistinction to the vibration number ν of the *forced* emission. If by (1a) we substitute $\frac{c\mathbf{E}}{\omega}$ for $\sin \omega t$, it follows from (17) that

$$M = - \frac{e^2}{2\pi^2 m} \sum_j \frac{A_{jk}q_{jk}}{\nu_{jk}^2 - \nu^2} \mathbf{E}.$$

This is the electric moment, calculated by wave-mechanics, which an individual atom acquires under the influence of the electric field \mathbf{E} . If

we multiply it by the number N of atoms in unit volume and divide by the intensity of field, we get from

$$M_1 = \frac{NM}{E},$$

the moment per unit volume referred to unit electric field:

$$M_1 = - \frac{e^2 N}{4\pi m} \sum_j \frac{2A_{jk}q_{jk}}{\nu_{jk}^2 - \nu^2} \quad (18)$$

This moment, multiplied by 4π , is according to the classical theory of electrons, equal to $n^2 - 1$, where n is the refractive index. We take this interpretation over into wave-mechanics. Thus we use (18) to *define* the refractive index n by means of the "dispersion formula"

$$n^2 - 1 = \frac{e^2 N}{\pi m} \sum_j \frac{2A_{jk}q_{jk}}{\nu_{jk}^2 - \nu^2} \quad (19)$$

The direction of q which has so far been left undetermined must be made to coincide with the direction of E (that is, the y -direction of our figure), as will be explained in the following section.

We may write (19) in a somewhat clearer form by again using the circular frequency ω and the corresponding circular frequency $\omega_{jk} = 2\pi\nu_{jk}$:

$$n^2 - 1 = \frac{4\pi e^2 N}{m} \sum_j \frac{2A_{jk}q_{jk}}{\omega_{jk}^2 - \omega^2} \quad (20)$$

This is precisely the classical dispersion formula both as regards the term before the summation sign and the general form of the "resonance-denominator." But there is the fundamental difference that *in the classical theory* we have in place of the transition frequencies ω_{jk} the proper frequencies themselves [which in our calculations were proportional to the E_j 's and not to the $(E_k - E_j)$'s]. It is very characteristic of the power of wave-mechanics that the introduction of perturbation theory in the integration of eqn. (8) makes the differences $E_k - E_j$ automatically replace the E_k 's that originally occurred in (8). The fact that experimentally only the transition frequencies ω_{jk} can be decisive in the dispersion formula scarcely needs mention in view of the character of all the measurements obtained in observations of anomalous dispersion.

C. Discussion of the Dispersion Formula. Calculation of the Number f of the Oscillators. The Summation Laws

Eqn. (20) contains a refinement which was first introduced into the classical dispersion formula by Kramers.* General interest, in view of the conditions of experiment, was concentrated on the dispersion of the atom in its *ground state*: the dispersion formula then contains only the

* H. A. Kramers, *Nature*, May and August, 1924; Kramers and Heisenberg, *Zeitschr. f. Phys.*, **31**, 684 (1925).

absorption frequencies of this ground state as the initial state (in the case of the alkalis, those frequencies are those of the principal series). In contrast with this Kramers considered the atom in an *arbitrary excited state* having the proper value E_k , and, using the correspondence-principle, he established the idea that in the dispersion formula all the *emission frequencies*, that are possible from the state E_k , must then be considered, that is, all the differences $E_k - E_j$, for which $E_j < E_k$. The wave-mechanical treatment confirms this point of view: as we emphasised in eqn. (11) the theory of perturbations compels us to take into account *all* states, those with $E_j < E_k$ as well as those with $E_j > E_k$.

We have next to consider the *numerator* of the dispersion formula. In the classical theory of electrons it denotes the *number of dispersion electrons* that participate in the proper vibration in question, or, expressed more generally, since this number, according to the quantum theory, need not be an integer, it denotes the "*number or quantity of the oscillators*"* (*Oszillatoren-Stärke*). By our formula (20) this number of the oscillators comes out as

$$f = -2A_{jk}q_{jk} \quad . \quad . \quad . \quad (21)$$

We have denoted it by f , as usual. It depends, on the one hand, on the probability of transition q_{jk} , that is, on the probability with which the transition $k \rightleftharpoons j$ would occur spontaneously, and on the other hand on the "probability of excitation" A_{jk} , that is, on the amount by which the partial vibration ψ_j in the perturbed state would be represented (resonance conditions not being considered) for a given state of polarisation of the perturbing field and for a given initial state ψ_k of the atom. Owing to the factor q_{jk} only such proper values E_j and corresponding frequencies ω_{jk} occur in the dispersion formula as can combine with the original proper value E_k . All transitions which are forbidden in emission also drop out in questions of dispersion, because for them we have, with $q_{jk} = 0$, also $f = 0$. For example, in the case of the alkalis in the ground state S only the lines of the principal series (SP) assert themselves in the dispersion formula, not the forbidden transitions (SD, SF) which contradict the rule of selection for the azimuthal quantum number. In § 7 E, Chap. I, we derived this rule of selection by showing, indeed, that, on account of the orthogonality of spherical harmonics, for these transitions we get $q_{jk} = 0$. It is true that in § 7 this proof was given explicitly only for the Coulombian field of the hydrogen atom, but it may be applied, § 8, Chap. I, to atoms with arbitrary central fields. In the same way we have calculated the dispersion explicitly only for the hydrogen atom; there is no difficulty in applying this proof to other atoms, so long as the mutual action between the electrons can be represented approximately by a central field.

* For further details on this point and concerning the relationship with Einstein's probabilities of transition, cf. R. Ladenburg, *Zeitschr. f. Phys.*, 7, 751 (1921).

we multiply it by the number N of atoms in unit volume and divide by the intensity of field, we get from

$$M_1 = \frac{NM}{E},$$

the moment per unit volume referred to unit electric field:

$$M_1 = - \frac{e^2 N}{4\pi m} \sum_j \frac{2A_{jk}q_{jk}}{\nu_{jk}^2 - \nu^2} \quad (18)$$

This moment, multiplied by 4π , is according to the classical theory of electrons, equal to $n^2 - 1$, where n is the refractive index. We take this interpretation over into wave-mechanics. Thus we use (18) to *define* the refractive index n by means of the "dispersion formula"

$$n^2 - 1 = \frac{e^2 N}{\pi m} \sum_j \frac{2A_{jk}q_{jk}}{\nu_{jk}^2 - \nu^2} \quad (19)$$

The direction of q which has so far been left undetermined must be made to coincide with the direction of E (that is, the y -direction of our figure), as will be explained in the following section.

We may write (19) in a somewhat clearer form by again using the circular frequency ω and the corresponding circular frequency $\omega_{jk} = 2\pi\nu_{jk}$:

$$n^2 - 1 = \frac{4\pi e^2 N}{m} \sum_j \frac{2A_{jk}q_{jk}}{\omega_{jk}^2 - \omega^2} \quad (20)$$

This is precisely the classical dispersion formula both as regards the term before the summation sign and the general form of the "resonance-denominator." But there is the fundamental difference that *in the classical theory* we have in place of the transition frequencies ω_{jk} the proper frequencies themselves [which in our calculations were proportional to the E_j 's and not to the $(E_k - E_j)$'s]. It is very characteristic of the power of wave-mechanics that the introduction of perturbation theory in the integration of eqn. (8) makes the differences $E_k - E_j$ automatically replace the E_k 's that originally occurred in (8). The fact that experimentally only the transition frequencies ω_{jk} can be decisive in the dispersion formula scarcely needs mention in view of the character of all the measurements obtained in observations of anomalous dispersion.

C. Discussion of the Dispersion Formula. Calculation of the Number f of the Oscillators. The Summation Laws

Eqn. (20) contains a refinement which was first introduced into the classical dispersion formula by Kramers.* General interest, in view of the conditions of experiment, was concentrated on the dispersion of the atom in its *ground state*: the dispersion formula then contains only the

* H. A. Kramers, *Nature*, May and August, 1924; Kramers and Heisenberg, *Zeitschr. f. Phys.*, **31**, 684 (1925).

We must here take S_{kj} as denoting the space factor S_{kj} (*Ortsfunktion*); we have already got rid of the factor denoting time dependence, $e^{\frac{2\pi i}{h}(E_k - E_j)t}$. But the current satisfies the relation

$$\operatorname{div} \mathbf{S} + \frac{\partial \rho}{\partial t} = 0.$$

If we write the last equation likewise without the time factors, then it runs in our case

$$\operatorname{div} S_{kj} + 2\pi i \nu_{kj} \rho_{kj} = 0; \quad \nu_{kj} = \frac{E_k - E_j}{h} \quad (26)$$

The integral on the right-hand side of (25) may now be written in a different form; we have

$$\int y \operatorname{div} S_{kj} d\tau = - \int S_{y, kj} d\tau, \quad (26a)$$

for the surface integrals again drop out in the integration by parts on account of the vanishing of S_{kj} at the "boundary." If we insert $\operatorname{div} \mathbf{S}$ from (26) in the left-hand side of (26a), we get

$$\int S_{y, kj} d\tau = 2\pi i \frac{E_k - E_j}{h} \int y \rho_{kj} d\tau.$$

Combined with (25), this gives

$$\left. \begin{aligned} A_{kj} &= - \frac{2\pi i m}{eh} \cdot \frac{2\pi i (E_k - E_j)}{h} \int y \rho_{kj} d\tau \\ &= \frac{4\pi^2 m (E_k - E_j)}{h^2} \int y \psi_k \psi_j^* d\tau = \frac{4\pi^2 m (E_k - E_j)}{h^2} \cdot y_{kj} \end{aligned} \right\} \quad (27)$$

This is the final result desired, namely, the coefficients A_{kj} are represented in terms of the co-ordinate matrix q_{kj} or its components y_{kj} . With the help of eqn. (27) we may also easily show that our formulæ (12) and (17) for the density and moment are identical with those of Schrödinger [Fourth Communication, eqns. (17) and (23)].

The oscillator quantity in eqn. (21) now become, if we insert the A_{kj} 's from (27),

$$f = - 2 \cdot \frac{4\pi^2 m}{h^2} (E_k - E_j) \cdot y_{kj} \cdot q_{kj} \quad (28)$$

thus—except for constant factors—the f 's become equal to the product of the proper frequency $\frac{E_k - E_j}{h}$ of the perturbed atom, of the co-ordinate-matrix element in the direction of polarisation of the incident light and of the co-ordinate matrix element in the direction of observation.

In eqn. (28), as we have remarked above, we must set $q_{kj} = y_{kj} z_n$. This is proved as follows. The way in which the proper functions depend

on the angle is the same in the case of an arbitrary atom with a central field as in the Kepler problem. It is given by

$$P_l^m(\cos \theta)e^{im\phi}.$$

If we take the y -axis as the polar axis $\theta = 0$, then $y = r \cos \theta$, $x \pm iz = r \sin \theta e^{\pm i\phi}$. For the moments y_{kj} and $(x \pm iz)_{kj}$ the same rules of selection apply as in the case of the Kepler problem, namely,

$$\begin{aligned} y_{kj} &= 0, \text{ except when } m = m', \\ (x \pm iz)_{kj} &= 0, \text{ except when } m = m' \pm 1. \end{aligned}$$

Here m or m' , respectively, is the magnetic quantum number that belongs to the state j or k . From this we see: if q is not chosen equal to y , then one of the two factors in the product $y_{kj} \cdot q_{kj}$ certainly vanishes. Consequently we must substitute $q_{kj} = y_{kj}$ in f , as was stated.

From eqn. (21) it is easy to prove the Summation Law of Thomas, Reiche, and Kuhn,* which was originally derived by these authors by means of the correspondence principle. In (21) we keep k constant and sum for f over all values of j , that is, over all transitions that lead to the level k . We then have

$$\sum_k f_{kj} = -2, \quad \sum_j A_{kj} q_{kj} = 1, \text{ if } q = y.$$

We have already mentioned on p. 32 that this summation law became of great importance for setting up the commutation law of quantum mechanics.

To prove it we form

$$\sum_j f_{kj} = -2 \sum_j A_{kj} \cdot q_{kj} = -2 \sum_j \int \psi_j \frac{\partial \psi_k}{\partial y} dt \int \psi_j q \psi_k \cdot dt \quad (\text{a})$$

in which we have used eqn. (22). The sum on the right-hand side is easy to evaluate. The $A_{kj} \cdot q_{kj}$'s can be represented as the expansion coefficients of two simple functions, as follows:

$$\frac{\partial \psi_k}{\partial y} = \sum_j A_{kj} \psi_j \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (\text{b})$$

because inversion gives

$$A_{kj} = \int \psi_j \frac{\partial \psi_k}{\partial y} dt.$$

Also,

$$q \psi_k = \sum_l q_{kl} \psi_l \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad (\text{c})$$

because it then follows that $q_{kl} = \int \psi_l q \psi_k$.

* W. Thomas, *Naturwiss.*, **13**, 627 (1925); W. Kuhn, *Zs. f. Phys.*, **33**, 403 (1925); F. Reiche and W. Thomas, *Zs. f. Phys.*, **34**, 510 (1925). Cf. also the analogous calculations in §7, eqns. (8) and (23).

If we now form the product of the functions on the left-hand sides of (b) and (c) and integrate over t , we get

$$\int \frac{\partial \psi_k}{\partial y} q \psi_k dt = \int \sum_j A_{kj} \cdot \psi_j \sum_l q_{kl} \psi_l dt = \sum_j A_{kj} \cdot q_{kj}$$

on account of orthogonality and the normalisation of the ψ 's. But the integral on the left further gives

$$\int \frac{\partial \psi_k}{\partial y} q \psi_k dt = \frac{1}{2} \int q \frac{\partial (\psi_k)^2}{\partial y} dt = -\frac{1}{2} \int \frac{\partial q}{\partial y} \psi_k^2 dt = \begin{cases} -\frac{1}{2}, & \text{if } q = y, \\ 0, & \text{if } q \neq y. \end{cases}$$

As above, we have here again omitted the surface integral. Substituting in (a) we get

$$\sum_j f_{kj} = \begin{cases} 1, & \text{if } q = y, \\ 0, & \text{if } q \neq y, \end{cases}$$

which completes the proof.

We must add that we have above proved the vanishing of f_{kj} for *central fields*, whereas the theorem $\sum f_{kj} = 0 (q \neq y)$ holds quite independently of the special assumption of a central field; in the proof we have actually made no such assumption. We further remark that in the case of n electrons 1 is to be replaced by n on the right-hand side of (d).

D. Non-coherent Scattering. Raman-Effect, Smekal Transitions

In § 5, Chap. I, we had to consider besides the density and moment of *one* atomic state also the corresponding quantities for the *transition* of the atom from *one* state to *another*, in order to be able to calculate the frequency and intensity of this transition, see eqns. (18), (19), and (21) of § 5, Chap. I [cf. also § 8, Chap. I, where the electric current was defined not only for a definite state but also for a *transition*]. We shall proceed similarly now with the atomic states disturbed by the light-wave. Thus we shall consider besides the state u of eqn. (11), which we allocate to the proper function ψ_k , for example, the fundamental vibration, also a second state which arises from the proper vibration ψ_l owing to the disturbing light-wave. The density combined from both states is then

$$u_k u_l^* = \left(\psi_k e^{\frac{2\pi i}{h} \mathbb{E}_k t} - \frac{hi}{4\pi c} \frac{e}{m} a \left\{ \begin{array}{c} \\ \end{array} \right\}_k \right) \left(\psi_l^* e^{-\frac{2\pi i}{h} \mathbb{E}_l t} + \frac{hi}{4\pi c} \frac{e}{m} a \left\{ \begin{array}{c} \\ \end{array} \right\}_l^* \right),$$

in which the abbreviations $\left\{ \begin{array}{c} \\ \end{array} \right\}_k$ and $\left\{ \begin{array}{c} \\ \end{array} \right\}_l^*$ are to be taken from (11), the first directly, the second after exchanging k for l and $+i$ for $-i$. Multiplying out and neglecting the term in a^2 gives us the following equation analogous to (12):

$$u_k u_l^* = \left. \begin{aligned} & \psi_k \psi_l^* e^{\frac{2\pi i}{h} (\mathbb{E}_k - \mathbb{E}_l) t} \\ & + \frac{hi}{4\pi c} \frac{e}{m} a \left(\psi_k \left\{ \begin{array}{c} \\ \end{array} \right\}_l^* e^{\frac{2\pi i}{h} \mathbb{E}_k t} - \psi_l^* \left\{ \begin{array}{c} \\ \end{array} \right\}_k e^{-\frac{2\pi i}{h} \mathbb{E}_l t} \right) \end{aligned} \right\} \quad (29)$$

The first term on the right-hand side expresses the time-dependence of the spontaneous transition $l \rightarrow k$, which does not interest us here. The time-dependence of the second member is given as we can see directly from (11) and (29), by the two factors

$$\frac{1}{e^{2\pi i(\frac{E_k - E_l}{h} + \nu)t}} \quad \text{and} \quad \frac{1}{e^{2\pi i(\frac{E_k - E_l}{h} - \nu)t}} \quad (30)$$

each of which occurs twice. Thus we have an accompanying vibration with *changed frequency*, a non-coherent scattering, besides the ordinary dispersion, which occurs with the frequency ν of the incident wave. Whereas the latter represents a *forced vibration* we can regard the former

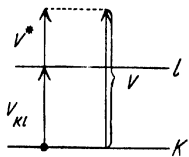


FIG. 17a.—The atom is in the ground state k . The incident frequency ν simultaneously effects the excitation of the atom (causing the transition $k \rightarrow l$) and the emission of light of the lesser frequency $\nu^* = \nu - \nu_{kl} < \nu$; in this process $h\nu_{kl} = E_l - E_k$.

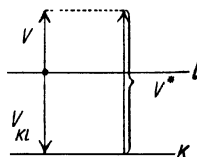


FIG. 17b.—The atom is in the excited state l . Under the influence of the incident frequency ν it passes into the state k of lesser energy and radiates light of higher frequency $\nu^* = \nu + \nu_{kl} > \nu$; here again we have $h\nu_{kl} = E_l - E_k$.

as *something between a free and a forced vibration*. If we signify the frequency due to the transition $l \rightarrow k$ by ν_{kl} , that is, if we set

$$\frac{E_l - E_k}{h} = \nu_{kl},$$

the vibration numbers of our non-coherent scattering are, according to (30), given by

$$\nu - \nu_{kl} \quad \text{and} \quad \nu + \nu_{kl}$$

respectively (the sign of i in the exponent is indeterminate).

The possible existence of this non-coherent scattering was proved by A. Smekal even before the advent of wave-mechanics on the basis of the following figures and formulæ.

The experimental proof of these "Smekal transitions" has very recently been achieved by C. V. Raman* and K. S. Krishnan. In light scattered by liquids, crystals, and gases,† they observed besides the ordinary scattered light of the same frequency also scattered light of less frequency than that of the incident beam, corresponding to Fig. 17a, and in some cases also light of higher frequency, corresponding to Fig. 17b.

* Indian Journ. of Physics, 2, March, 1928; cf. also Nature, 121, 501, and 122, 12 (1928).

† Raman and Krishnan have obtained particularly sharp modified lines with crystals. Such lines were also found for quartz and calcite by P. Pringsheim and B. Rosen (Zeitschr. f. Phys., 50, 741 (1928)); R. W. Wood (Phil. Mag., Ser. 7, 6 (1928)); and by G. Landsberg and L. Mandelstam (Naturwiss., 16, 568 (1928); Comptes Rendus, 187, 109 (1928)). The last two investigators performed their researches independently of Raman and Krishnan and about the same time. They explained their discovery in the same way as the latter.

It is self-evident that the phenomenon depicted in Fig. 17b will occur with considerably less intensity than that represented in Fig. 17a, because it assumes the presence of excited atoms. This is in agreement with the findings of the Indian physicists quoted.

This discovery is the optical analogy to the Compton effect, as will be clearly evident in an account in § 7. This characterises the fundamental importance of the Raman effect, which is further emphasised by the fact that the displacements ν_{kl} of the primary radiation can be identified in many cases* with infra-red vibration-frequencies of the molecules investigated (for example, benzol, toluol, water).

In his wave-mechanical representation of the Smekal leaps Schrödinger assumes, as in his derivation of the ordinary Bohr frequency condition (cf. p. 45), that the states k and l that here come into question are actually excited in the atom. We have avoided this assumption, which is probably too special and too closely bound up with the model idea, and we have contented ourselves with generalising the dispersion theory formally by considering (as formerly in the case of the frequency condition) the density matrix $\rho_{kl} = u_k u_l^*$ for the two states combined.

We shall proceed in the same way in making statements about the intensities of Smekal transitions. For this purpose we shall have to consider the "moment-matrix" M_{kl} of eqn. (19), § 5, Chap. I. We shall work out the formulæ only as far as will be required later for the theory of the Compton effect. With this in view we shall distinguish between A_j^+ and A_j^- in eqn. (9); A_j^+ is to belong to w_+ , that is to the upper (negative) sign of the exponent in (9), and A_j^- to the lower. Further, for the sake of generality and on account of its applicability to the Compton effect we shall assume ψ complex. The A 's must be furnished with double indices, as in eqns. (17) to (20), so we write them A_{jk} and A_{jl} according as they belong to $\partial\psi_k/\partial y$ or $\partial\psi_l/\partial y$ [cf. eqn. (9)].

We next form the factors of the exponential quantities (30) in the expression (29) for $u_k u_l^*$, and obtain respectively, as in eqns. (12) and (12a):

$$\frac{hi}{4\pi c} \frac{e}{m} \sum \pm, \quad \sum \pm = \sum_j \left(\frac{A_{jl}^* \psi_k \psi_j^*}{E_l - E_j \mp h\nu} - \frac{A_{jk}^* \psi_l^* \psi_j}{E_k - E_j \pm h\nu} \right) \quad (31)$$

Passing on to the "moment-matrix" M_{kl} , we have to form the moment of the "electric density" $eu_k u_l^*$ in any arbitrary co-ordinate direction q , that is, corresponding to eqn. (19) of § 5, Chap. I:

$$M_{kl} = e \int q u_k u_l^* d\tau \quad (32)$$

Integration with respect to $d\tau$ over the co-ordinate space is only to be

* The fact that they do not in all cases correspond with infra-red absorption lines is due to the selection rules in the Raman effect being different from those in the oscillation spectra. The Raman lines occur when there is a level j which can combine with k and l , but the infra-red lines can occur only if k and l can combine directly.

for the first time, indeed, the directional distribution of photo-electrons can be calculated by means of wave-mechanics.* We must admit, however, that Einstein's law, like the Bohr frequency condition which is related to it, is not actually derived, but is included in the basic assumptions of wave-mechanics (cf. Chap. I, p. 46).

The wave-mechanical theory of the photo-electric effect links up closely with the preceding section. We need only direct our attention to the *continuous* spectrum instead of to the *line* spectrum and apply to this continuous spectrum the calculus of perturbations. We shall take this opportunity of adding that in the formulæ of the preceding section we have also to imagine the frequencies of the *continuous* spectrum to be included, that is, that (cf. the end of § 1 A) to the summation over the discrete spectrum must be added the summation over the continuous spectrum, just as will be done in what immediately follows.

A. Perturbation in the Continuous Spectrum

We shall here consider not the usual method of producing the photo-electric effect, in which light of short wave-length falls on the surface of a metal and forces the *free* electrons of the metal to be emitted, but the photo-electric effect in the individual atom, in particular the hydrogen atom, in which case, therefore, the light acts on a *bound* electron. This problem is more definite theoretically than the ordinary photo-electric effect and has also received strong experimental support by modern methods (Wilson's cloud-chamber).

We describe the incident wave of light as in § 3 by the assumption (1, 1a) and by Fig. 16. The perturbed state of the atom it produces is given by eqn. (11), § 3. But whereas earlier we wrote down the sums only for the discrete proper values we shall now express only the integral over the continuous spectrum of proper values and shall denote the additional sum by. . . Let the proper value parameter of the continuous spectrum be E' , and the proper function $\psi(E')$. Further, let E_k and ψ_k denote as earlier the initial state, that is, in particular, the ground state of the atom. Eqn. (11), § 3, then runs (μ = mass of the electron)

$$u = \psi_k e^{\frac{2\pi i}{h} E_k t} + \dots - \frac{hi}{4\pi c} \frac{e}{\mu} a\{I + II\} \quad (1)$$

where

$$I = \int \frac{A(E')\psi(E')dE'}{E_k - E' + h\nu} e^{\frac{2\pi i}{h}(E_k + h\nu)t},$$

$$II = \int \frac{A(E')\psi(E')dE'}{E_k - E' - h\nu} e^{\frac{2\pi i}{h}(E_k - h\nu)t}.$$

The earlier quantity A_j is now replaced by the quantity $A(E')dE'$.

* G. Wentzel, *Zeitschr. f. Phys.*, **40**, 574 (1926); **41**, 828 (1927); G. Beck, *ibid.*, **41**, 443 (1927). We shall go beyond the results of these authors, particularly in the treatment of short-wave radiation.

Corresponding to eqn. (9), § 3, it is defined by

$$\frac{\partial \psi}{\partial y} e^{\mp 2\pi i \frac{x}{\lambda}} = \dots + \int A(E') \psi(E') dE' \quad (2)$$

λ is the wave-length of the incident light.

The integrals in (1) and (2) are to be taken over the region $E' = E_0 = \mu c^2$ to $E' = \infty$. The lower limit $E' = E_0$ corresponds in the normalisation of § 7, Chap. I, to the value $E' = 0$ or, expressed in the language of orbits, to removing the electron from the atom to infinity, so that $E' > E_0$ is equivalent to the condition in § 7: $E > 0$ for the continuous spectrum. (The . . . in (2) denote that, as in every representation of a function in terms of proper functions the discrete *and* the continuous proper values are to be taken into account, but that we have written down only the continuous values.)

We now consider a single element of the integral I (we shall return to II in Note 3, p. 293) and separate it into three factors:

$$(a) \frac{A(E') dE}{E_k - E' + h\nu}, \quad (b) \psi(E'), \quad (c) e^{\frac{2\pi i}{h}(E_k + h\nu)t} \quad (3)$$

(a) is constant, that is, independent of the space and time co-ordinates and dependent only on the quantum numbers; (b) signifies a distribution in space of the proper vibration; (c) denotes its dependence on time.

Let us first deal with (b). We examined the proper functions of the hydrogen atom in the region of the continuous spectrum in § 7 A, (b), of Chap. I. We shall give a more detailed treatment of them in Note I at the end of the book. We shall there split up the single proper function $\psi(E')$ into two parts ψ^1 and ψ^2 , so that

$$\psi = \frac{1}{2}(\psi^1 + \psi^2).$$

The asymptotic expression in eqn. (24) of § 7, Chap. I, belongs to the radial component R of these two parts; the sign + in the exponent may be attributed to ψ^2 , the sign - to ψ^1 . To this radial component must be added the angular component from the eqn. (2) of the same section; altogether we get asymptotically

$$\frac{\psi^1(E')}{\psi^2(E')} \Big\} = \frac{C}{r} e^{\pm i(kr + a)} P_l^m(\cos \theta) e^{im\phi} \quad (4)$$

The radial component represents a spherical wave, which is radiated out from $r = 0$ or which with the sign reversed converges to $r = 0$. Just as in § 1, Chap. I, we regarded the plane de Broglie wave e^{ikx} as representing a parallel stream of electrons, so we shall have to interpret the spherical wave as representing a *stream of electrons diverging from or converging to the atom*. This interpretation is justified by the following characteristic difference between the asymptotic behaviour of the continuous proper functions and that of the discrete proper functions: the discrete proper functions ψ vanish so rapidly at infinity that the integral $\int \psi^2 d\sigma$ taken over an infinitely great sphere vanishes; for the con-

tinuous proper functions, on the other hand, this integral has a finite value independent of r . Hence the former proper functions indicate the charge zero at infinity and correspond to the *elliptic orbits* formerly used; whereas the latter proper functions indicate a charge escaping to infinity, and correspond to the classical *hyperbolic paths*. Thus we have a *stream of electrons* diverging from or converging to the atom, the statistical behaviour of which is described by the wave-function. The different directions are not equivalent but are distinguished from each other by the dependence on θ and ϕ contained in (4).

With our choice of sign in the time factor $\left(e^{\frac{2\pi i}{h}(E_k + h\nu)t}\right)$ we see that $\psi^2 \cdot e^{-i(kr + a)}$ gives a spherical wave emitted *outwards* by the atom, but $\psi^1 \cdot e^{+i(kr + a)}$ gives a spherical wave radiating *inwards* towards the atom from infinity. Only the first result has a physical meaning. In Note 3 we shall show that in carrying out the integration expressed in (1) the partial function ψ^1 drops out and only ψ^2 stands. An essential feature is the choice of the path of integration which will also be discussed in Note 3.

On the basis of these remarks we shall in the sequel replace the factor (b), which gives the distribution of the photo-electrons in space by $\frac{1}{2}\psi^2$. On the other hand, we must use the proper function ψ and not the partial function $\frac{1}{2}\psi^2$ in the factor (a), namely in the definition of $A(E')$, eqns. (2) and (7).

We next turn to the time-factor (c) in (3). Just as in the case of the plane de Broglie wave, the factor $2\pi i/\hbar$ gives us the total energy of the electron emitted. Thus it is

$$E_k + h\nu.$$

If we deduct from this the proper-energy $E_0 = m_0c^2$ corresponding to the electron at rest, we get

$$\epsilon = E_k - E_0 + h\nu \quad . \quad . \quad . \quad . \quad (5)$$

the kinetic energy of the emitted electron. But $E_0 - E_k$ is the ionisation potential J , measured in terms of energy, or, what amounts to the same thing, the binding energy of the electron in the atomic state E_k compared with its state of freedom from the atom E_0 . Thus eqn. (5) is identical with

$$\epsilon = h\nu - J \quad . \quad . \quad . \quad . \quad (6)$$

and this is identically *Einstein's photo-electric equation*. Although not explained or made intelligible here, it is at least closely interwoven with the fundamental assumptions of wave-mechanics. It does receive some explanation, on the other hand, from the conception of discrete energy-elements of magnitude $h\nu$ contained in the light, that is, from the well-known hypothesis of light-quanta.

B. The Angular Distribution of Photo-Electrons for the Case of Long-Wave Light

This angular distribution is indicated generally by the factor $P_l^m(\cos \theta) e^{im\phi}$. For arbitrary values of the quantum numbers l and m it includes many different possibilities (subdivisions with respect to the angle θ as well as with respect to ϕ). But we shall show that for the choice of co-ordinates to be given in (9) these many different possibilities are restricted by the factor (3a) to the quantum numbers $l = 1$ and $m = 0$. The factor (3a) contains the quantity $A(E')$ which is determined by eqn. (2). If we first consider only long wave light (λ very large compared with atomic dimensions), we can then set the exponential factor in the left-hand side of (2) equal to unity, as we have done in the preceding section. The inverse (*Umkehrung*) of (2) then gives us according to the general rules for normalised proper functions [cf. in the preceding section the inverse of eqn. (9), which we actually gave in eqn. (22)]:

$$A(E') = \int \frac{\partial \psi_k}{\partial y} \psi_k^*(E') d\tau \quad . \quad . \quad . \quad (7)$$

The integration is to be taken over the whole of the co-ordinate space, and if polar co-ordinates are used we have

$$d\tau = r^2 dr \sin \theta d\theta d\phi \quad . \quad . \quad . \quad (7a)$$

We take ψ_k as the fundamental solution for the hydrogen atom, thus

$$\psi_k = Ce^{-Zr/a},$$

where C is the normalising factor and a the radius of the first Bohr circle (cf. the table on p. 71 and eqn. (30a) on p. 70; in the case of hydrogen $Z = 1$). Hence

$$\frac{\partial \psi_k}{\partial y} = -\frac{Z}{a} Ce^{-Zr/a} \cdot \frac{y}{r} \quad . \quad . \quad . \quad (8)$$

We define our polar co-ordinates in such a way that the y -axis (direction of the electric force of the light-wave) is the polar axis. Thus we set

$$y = r \cos \theta, \quad x = r \sin \theta \cos \phi, \quad z = r \sin \theta \sin \phi \quad . \quad (9)$$

We then get $\partial \psi_k / \partial y$ proportional to $\cos \theta = P_1(\cos \theta)$, and by (4), (7), (7a) and (8) we get

$$A(E') = . . . \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi P_1(\cos \theta) P_l^m(\cos \theta) e^{-im\phi} \quad . \quad (10)$$

The dots . . . here denote the radial part of the integral, which does not at present interest us. From (10) we at once infer that $A(E')$ differs from zero only if

$$m = 0 \quad \text{and} \quad l = 1 \quad \text{simultaneously} \quad . \quad . \quad . \quad (11)$$

The general way in which $\psi(E')$ depends on the direction, $P_l^m(\cos \theta) e^{im\phi}$, thus reduces to

$$P_1(\cos \theta) = \cos \theta, \text{ which is independent of } \phi. \quad (12)$$

We must regard the square of ψ as giving the *probable number of the photo-electrons* for every direction θ (in other words, this gives the intensity J emitted in the spherical wave in the direction in question). Thus this number, or intensity, is given by

$$J \propto \cos^2 \theta \quad (13)$$

and is represented in Fig. 18. In accordance with the meaning of θ we must, of course, picture Fig. 18 as rotated about the y -axis and repeated in the negative direction of y

$$\left(\frac{\pi}{2} \leq \theta \leq \pi\right),$$

so that there are two pear-shaped symmetrical bodies of rotation in space.

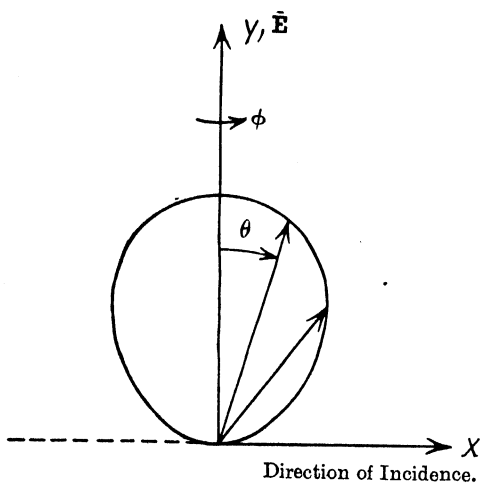


FIG. 18.—Photo-electric emission in the case of long waves.

The intensity is plotted in polar co-ordinates.

To make the conclusion which brought us to eqn. (11) more convincing analytically we make the following remark. Just as in the preceding section we had to take the summation sign, for example, in the dispersion formula as applying to *all* quantum numbers, so here we must regard the process indicated by our integration signs as applying to *all* quantum numbers, summation being performed over the discrete values and integration over the *continuous* values. Thus we must form the sum for

the angular quantum numbers l and m , and we must integrate only over the radial quantum number or the equivalent principal quantum number (or total energy E'). So the integral sign in eqn. (1) when written out in full, stands for

$$\sum_l \sum_m \int dE' \quad (14)$$

and the function $A(E')$ should really be replaced by

$$A_{lm}(E').$$

What we have shown above amounts to proving that of the doubly infinite sum (14) of integrals only one term differs from zero, namely that for which $A_{lm} = A_{10}$, whereas all other A_{lm} 's vanish.

If we look at Fig. 18 the impression forces itself upon us that the electric force by its rapidly alternating effects on the atom releases the photo-electron from the atomic configuration and impresses its own direction on it for preference; further, that deviations from this favoured direction are in some way due to the initial state of the electron. It is very remarkable that wave-mechanics gives no further details of the causal mechanism sketched in this way, but proceeds to deal straightway with the statistical final effect, which alone comes into consideration for comparison with experiment. Whereas the light-quantum view of the Einstein law appears to depict the *causal aspect* of the phenomenon convincingly, wave-mechanics abandons causal arguments and, instead, gives full expression to the *statistical aspect*. The statistical character of the method of wave-mechanics also manifests itself in the fact that even in the case of hydrogen with its single electron, the theory leads to a continuous curve of charge-distribution which is obviously to be regarded as the mean of a very great number of individual cases.

To supplement the preceding calculations and to prepare for the next section we shall make a second calculation in which we take our polar axis $\theta = 0$ not in the direction of the electric force (y -axis) but in the direction of emission (the x -axis), by setting (cf. Fig. 19, p. 187)

$$x = r \cos \theta, \quad y = r \sin \theta \cos \phi, \quad z = r \sin \theta \sin \phi \quad (9a)$$

Then

$$\frac{\partial \psi_k}{\partial y} = -\frac{Z}{a} C e^{-Zr/a} \sin \theta \cos \phi \quad (8a)$$

for which, by eqn. (12), § 2, Chap. I, and suppressing the radial component, we may set

$$\frac{\partial \psi_k}{\partial y} = \dots P_1^1(\cos \theta) \cos \phi \quad (8b)$$

Eqn. (10) then becomes

$$A(E') = \dots \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} d\phi P_1^1(\cos \theta) \cos \phi P_l^m(\cos \theta) e^{-im\phi} \quad (10a)$$

From this it follows that $|m| = 1$ and that $l = 1$; the latter on account of the condition of orthogonality of the associated spherical harmonics, eqn. (24), § 6, Chap. I. Hence the dependence of $\psi(E')$ on the angle now reduces to $P_1^1(\cos \theta) e^{\pm i\phi} = \sin \theta e^{\pm i\phi}$ or rather to

$$\psi(E') = \sin \theta \cos \phi \quad (12a)$$

since of the two possibilities $\cos \phi$ and $\sin \phi$ included in $e^{\pm i\phi}$, only the first leads to a value of $A(E')$ different from zero, whereas the second is

again suppressed owing to the vanishing of the factor A . In accordance with this we have in place of (13)

$$J \sim \sin^2 \theta \cos^2 \phi \quad . \quad . \quad . \quad (13a)$$

The direction distribution given by this expression is of course again represented by Fig. 18: in view of the new position of the polar axis the maximum now lies at $\theta = \pi/2$, $\phi = 0$; the rotational symmetry about the direction of the electric force is not so immediately evident in the new representation as in the old. But the fact that both representations are identical can be seen directly in rectilinear co-ordinates: the two expressions (13) and (13a) both give by (9) and (9a):

$$J \sim \frac{y^2}{r^2}.$$

Here y and r at first denote the co-ordinates of the point at which the intensity J is observed. But we may also regard them as the co-ordinates of the end-point of J in a polar diagram which, as in Fig. 18, represents the magnitude of J for every direction by means of the radius vector r . The equation of this polar diagram then runs in both cases simply

$$r^3 = y^2,$$

which expresses besides the identity of both methods of representation also the rotational symmetry about the y -axis.

The total number of photo-electrons $J_\theta d\theta$ which is emitted in a cone between θ and $\theta + d\theta$ is obtained from (13a) by multiplying with $\sin \theta d\theta d\phi$ and integrating over ϕ . In this way we get

$$J_\theta \sim \sin^3 \theta \quad . \quad . \quad . \quad (13b)$$

C. The One-sidedness of Electronic Emission in the case of Light of Short Wave-Length

The whole difference between this case and the case of light of long wave-length consists in the fact that the exponential factor in (2) is not neglected now. We take the negative sign in the exponential factor, because it belongs to the integral I of eqn. (1) and because II is of no direct physical interest (cf. Note 3). By analogy with (7) we then get

$$A(E') = \int \frac{\partial \psi_k}{\partial y} e^{-2\pi i \frac{r}{\lambda}} \psi^*(E') d\tau.$$

For convenience we use the polar co-ordinates (9a) and, in view of (4) and (8a) we get

$$A(E') = -\frac{Z}{a} C \int_0^\infty e^{-\frac{Zr}{a}} R_l r^2 dr A_{lm} \quad . \quad . \quad (15)$$

$$A_{lm} = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \sin \theta \cos \phi e^{-2\pi i \frac{r}{\lambda} \cos \theta} P_l^m(\cos \theta) e^{-im\phi} \quad (15a)$$

In (15) R_l denotes the radial component of $\psi^*(E')$; the suffix l indicates that it will in general depend on l . We shall imagine the normalising factor of the proper function included in it.

From (15a) it necessarily follows that $|m| = 1$, as A vanishes for $m = 0$ and $|m| > 1$. But A would also vanish for $|m| = 1$, if we were to substitute $\sin \phi$ for $e^{-i\phi}$ [cf. what was said in relation to eqn. (12a)]. The only value of A_{lm} which does not vanish is therefore

$$A_{11} = \pi \int_0^\pi \sin^2 \theta e^{-2\pi i \frac{r}{\lambda} \cos \theta} P_1^1(\cos \theta) d\theta \quad . \quad . \quad (16)$$

We next expand the exponential function in (16) and retain only the 0th and 1st member of the power series, by assuming that λ is still large compared with the atomic dimensions r that come into question. We thus obtain

$$A_{11} = \pi \int_0^\pi \sin^2 \theta \left(1 - \frac{2\pi i r}{\lambda} \cos \theta \right) P_1^1(\cos \theta) d\theta \quad . \quad . \quad (17)$$

The "0th" term of course agrees with the integral (10a), since this denoted the approximation for an infinitely great λ , provided we set $m=1$ in it and integrate over ϕ . As in the case of (10a) it follows that this term differs from zero only for $l = 1$. We denote its value for $l = 1$ by A_0 :

$$A_0 = \pi \int_0^\pi \sin^2 \theta d\theta = \pi \int_{-1}^{+1} (1 - x^2) dx = \frac{4}{3}\pi \quad . \quad . \quad (17a)$$

In the second term of (17) we introduce the abbreviation A_1 by writing

$$A_1 = \pi \int_0^\pi \sin^2 \theta \cos \theta P_1^1(\cos \theta) d\theta \quad . \quad . \quad (17b)$$

We calculate A_1 by making use of the expression for P_l^m in eqn. (12) on p. 12:

$$A_1 = \pi \int_0^\pi \sin^2 \theta \cos \theta \frac{dP_1(\cos \theta)}{d \cos \theta} d\theta = \pi \int_{-1}^{+1} x(1 - x^2) \frac{dP_1(x)}{dx} dx \quad (17c)$$

or by integrating by parts once:

$$A_1 = \pi \int_{-1}^{+1} (3x^2 - 1) P_1(x) dx.$$

But by Fig. (1) p. 12, the first factor under the integral sign is equal to $2P_2(x)$. Thus

$$A_1 = 2\pi \int_{-1}^{+1} P_2(x) P_1(x) dx.$$

186 Chapter II. Perturbations. Diffraction Problems

From this we infer: A_1 differs from zero only for $l = 2$, and by eqn. (30) on p. 54 it becomes in this case equal to

$$2\pi \cdot \frac{2}{2l+1} = \frac{4\pi}{5}.$$

Recapitulating, we see that the sum denoted in (14) reduces in our choice of co-ordinate system and if we neglect higher powers of r/λ to two terms. For the one we have $m = 1$ and $l = 1$; for the other, $m = 1$ and $l = 2$:

1. $m = 1, l = 1$. By (15) we have

$$A(E') = -\frac{Z}{a} C \cdot K_1 A_{11}, \quad K_1 = \int_0^\infty e^{-\frac{Zr}{a}} R_1 r^2 dr.$$

2. $m = 1, l = 2$. By (15) we now have

$$A(E') = \frac{2\pi i}{\lambda} \frac{Z}{a} C K_2 A_{21}; \quad K_2 = \int_0^\infty e^{-\frac{Zr}{a}} R_2 r^3 dr.$$

The coefficients A_{11} and A_{21} that here occur are obviously identical with the abbreviations A_0, A_1 of eqns. (17a, b). We have to multiply these A 's with the correspondingly defined partial functions ψ^2 in note 1 (and *not*, cf. p. 179, with the proper functions ψ themselves); these ψ^2 's are to be normalised exactly like the ψ 's. These partial functions are:

$$\begin{aligned} 1. \quad m = 1, l = 1 \\ \psi^2 &= R_1^2 P_1^1 (\cos \theta) \cos \phi. \\ 2. \quad m = 1, l = 2. \\ \psi^2 &= R_2^2 P_2^1 (\cos \theta) \cos \phi. \end{aligned}$$

Inserting the numerical values of the A 's and also the expressions P_1^1 and P_2^1 we get for the angle dependence of the sum indicated in (14)

$$\sum \sum A(E') \psi(E') = \dots \left\{ 1 - \frac{18}{5} \frac{\pi i}{\lambda} \frac{K_2}{K_1} \frac{R_2^2}{R_1^2} \cos \theta \right\} \sin \theta \cos \phi \quad (18)$$

Here we have placed the factor of the first term independent of θ, ϕ before the bracket.

We shall occupy ourselves with the evaluation of the integrals K_1 and K_2 in Note 2. We insert in (18) the values there obtained as well as the asymptotic values of R_2^2 and R_1^2 calculated in Note 1.

If we denote the left-hand side of (18) by Ψ , we have

$$\Psi \sim \left\{ 1 + \frac{4\pi}{\lambda k} \frac{2 + iZ/ak}{1 + Z^2/a^2 k^2} \cos \theta \right\} \sin \theta \cos \phi.$$

The square of the absolute value of Ψ now takes the place of the earlier J , eqn. (13) or (13a), and gives the measure of the direction of distribution

of the photo-electrons. By neglecting λ^{-2} , which is justifiable, we get

$$J = |\Psi|^2 \sim \left\{ 1 + \frac{16\pi}{\lambda} \frac{k \cos \theta}{k^2 + Z^2/a^2} \right\} \sin^2 \theta \cos^2 \phi. \quad (19)$$

The dependence on the azimuth ϕ remains, of course, the same as before, but the dependence on θ has changed.

We shall show that in consequence of (19) the *emission of electrons is favoured in the forward direction*, that is, that the distribution of electrons inclines towards the forward direction of the rays. For this purpose we calculate the solid angle of the "bisecting cone" (*Halbierungskegel*), that is, of the cone that includes half of all the emitted electrons. Let the θ -value of the bisecting cone be called Θ , then

$$\int_0^{2\pi} \int_0^\Theta J \sin \theta \, d\theta \, d\phi = \int_0^{2\pi} \int_\Theta^\pi J \sin \theta \, d\theta \, d\phi.$$

Integration with respect to ϕ leads to the same constant factor on both sides, which we may neglect; the integration with respect to θ is

$$\int_0^\Theta \left(1 + \frac{16\pi}{\lambda} \frac{k \cos \theta}{k^2 + Z^2/a^2} \right) \sin^3 \theta \, d\theta = \int_\Theta^\pi \left(1 + \frac{16\pi}{\lambda} \frac{k \cos \theta}{k^2 + Z^2/a^2} \right) \sin^3 \theta \, d\theta.$$

We set $x = \cos \theta$, $x_1 = \cos \Theta$ and

$$\frac{16\pi}{\lambda} \frac{k}{k^2 + Z^2/a^2} = \kappa \quad (20)$$

Θ does not differ much from $\pi/2$, for with our assumptions (the wave-length λ of the incident light was still to be sufficiently large to allow powers of r/λ higher than the first to be neglected), the distribution will still be very similar to that which occurs in the case of light of long wave-length, which was symmetrical to $\theta = \pi/2$; for this reason x_1 becomes small compared with 1. On integrating we obtain, if we strike out higher powers of x_1 ,

$$x_1 = \frac{\kappa}{4}.$$

If we set

$$\Theta = \frac{\pi}{2} - \theta_0,$$

then, for our approximation $x_1 = \cos \Theta = \theta_0$,

$$\theta_0 = \frac{\kappa}{4} = \frac{4\pi}{\lambda} \frac{k}{k^2 + Z^2/a^2} \quad (21)$$

thus gives us the deviation of the aperture of the bisecting cone from $\pi/2$. Since θ_0 is positive the electrons show a preference for the *forward direction*, as we have asserted.

The position of the maximum of photo-electric emission has also

changed, being likewise displaced forwards of the $\pi/2$ direction. We calculate it by (19) and (20) from the equation

$$0 = \frac{d}{d\theta} \sin^2 \theta \{1 + \kappa \cos \theta\} = 2 \sin \theta \left(\cos \theta + \kappa \cos^2 \theta - \frac{\kappa}{2} \sin^2 \theta \right).$$

In the last bracket we may simplify the terms multiplied by the small quantity κ by substituting for them as a first approximation $\theta = \pi/2$. The bracketed expression then gives for the locus of the maximum

$$\left. \begin{aligned} \cos \theta &= \frac{\kappa}{2} > 0, \\ \theta &= \theta_{\max} \\ &= \frac{\pi}{2} - \frac{\kappa}{2} \end{aligned} \right\} \quad \dots \quad (22)$$

Thus the maximum has likewise become displaced in the forward direction; the amount of the displacement is twice as great as in the case of the bisecting cone (cf. Fig. 19).

This position of θ_{\max} allows us to write the distribution law (19) for the region in the neighbourhood of the maximum more simply as follows:

$$J = \sin^2 \left(\theta + \frac{\kappa}{2} \right) \cos^2 \phi \quad \dots \quad (23)$$

For, from (23) it follows by expanding in terms of κ ,

$$J = \left(\sin \theta + \frac{\kappa}{2} \cos \theta \right)^2 \cos^2 \phi = (1 + \kappa \cot \theta) \sin^2 \theta \cos^2 \phi$$

and for the immediate neighbourhood of the maximum ($\sin \theta \sim 1$,

$\cot \theta \sim \cos \theta$) this becomes identical with (19). Eqn. (23) admits the following interpretation. The distribution of the photo-electrons in the case of light of short wave-length about the displaced maximum $\theta = \pi/2 - \kappa/2$ is in its immediate neighbourhood the same as for light of long wave-length about the undisplaced maximum $\theta = \pi/2$.

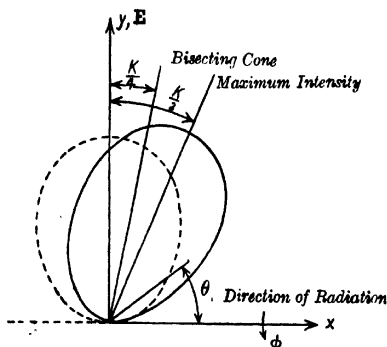


FIG. 19.—Photo-electric emission in the case of short waves (Röntgen rays). More intense emission forwards than backwards.

We shall reduce the displacement of the bisecting cone, eqn. (21), to a simpler form.

By eqns. (24a), (19), and (3a) of § 7, Chap. I, the meaning of k^2 was

$$k^2 = \frac{1}{r_0^2} = \frac{8\pi^2\mu}{h^2}(E' - E_0). \quad (24)$$

Corresponding to our present nomenclature and normalisation we have written $E' - E_0$ for the energy previously denoted by E ($E > 0$ in the continuous spectrum). Our value (24) thus depends on the particular point of the continuous spectrum that we happen to be considering. Now, we see in Note 3 that only the point $E' = E_k + h\nu$ as given by the resonance denominator in integral I, eqn. (1) is significant. If we insert its E' in (24) and take into account eqn. (5) in which $\epsilon = \mu v^2/2$ denoted the kinetic energy of the emitted electron, (24) yields simply

$$k = \frac{2\pi\mu}{h}v \quad (25)$$

(We may remark in passing that this may also be written

$$k = \frac{2\pi}{\lambda_v},$$

where λ_v denotes the de Broglie wave-length associated with v .) We next calculate the denominator in (21), namely $k^2 + Z^2/a^2$. If we insert the value (24) for k^2 and the value $E_k + h\nu$ for E' ,

$$k^2 + \frac{Z^2}{a^2} = \frac{8\pi^2\mu\nu}{h} + \frac{8\pi^2\mu}{h^2}(E_k - E_0) + \frac{Z^2}{a^2}.$$

On the right, however, the last two terms cancel out, for

$$a = \frac{h^2}{4\pi^2\mu e^2}, \quad \frac{Z^2}{a^2} = \frac{8\pi^2\mu}{h^2}RhZ^2$$

and for the ground state of the hydrogen atom with the nuclear charge Z

$$E_k - E_0 = -RhZ^2 \quad (R = \text{Rydberg constant}).$$

Accordingly there remains

$$k^2 + \frac{Z^2}{a^2} = \frac{8\pi^2\mu\nu}{h}. \quad (26)$$

If we substitute (25) and (26) in (21), it follows that

$$\frac{\kappa}{4} = \frac{v}{\lambda\nu} = \frac{v}{c}. \quad (27)$$

It suggests itself to us to bring this value into relationship with the momentum of the light-quantum which led to the emission of the electron. This momentum is $h\nu/c$ and has the direction of the ray. The momentum of the electron is μv and favours the direction perpendicular to the ray, except for deflection due to the momentum of the light-quantum. Thus the deflection amounts to

$$\frac{h\nu}{c} / \mu v \quad (28)$$

If we then use Einstein's law and, neglecting the ionisation potential J , set

$$h\nu = \frac{\mu v^2}{2},$$

(28) becomes identical with (27) except for the factor 2 to which we shall revert later.

When we just now spoke of the momentum of the light-quantum, we forsook the statistical basis of wave-mechanics for the *causal* basis of the light-quantum hypothesis which underlies Einstein's law. Instead of the momentum of the light-quantum we may also adopt a more phenomenological attitude and regard the *radiation pressure* of the light as the reason for the forward deflection of the photo-electrons, for this pressure is only another way of expressing momentum of the light-quantum.

But it must be added that in this way we arrive only at a qualitative explanation of the forward deflection of the photo-electrons; for a quantitative explanation we require wave-mechanics and the factor 2 which was emphasised just above.

Our treatment was restricted to the ground state of the hydrogen atom or, as we may say more generally, to the K-shell. We perceive immediately that the results for the L- and M-shells will come out differently on account of the proper functions in general depending differently on the angle of deflection.

Compared with the psychological advantages of the causal view the statistical wave-mechanical treatment here has the practical advantage, which we must again emphasise, of answering questions of intensity. We have seen this in calculating the relative distribution of photo-electrons. But we may go further and ask what is the total number of photo-electrons released, that is, the *absolute value of the photo-electric emission*? This would lead us to a rational theory of the absorption coefficient of light and Röntgen radiation; in the case of Röntgen radiation we should be concerned not with *one* absorption coefficient but with a series of such coefficients for the absorption phenomena in the K-shell, in the three L-shells, and so forth. Wentzel (*loc. cit.*) has sketched out this theory. Here we may point out that, besides being essential for understanding light and Röntgen phenomena this theory is also of fundamental importance for gaining a knowledge of the internal constitution of the stars (Eddington).

D. Comparison with Experiment

We shall first speak of the general character of the distribution of velocities which we studied in B, and shall then pass on to discuss a little more closely the one-sidedness of the distribution which was derived in C. Indisputable results can be obtained only in the case of light of sufficiently short wave-length, that is, Röntgen rays.

Eqn. (13), or (13a), refers to polarised rays. The initial direction of photo-electrons that are released by Röntgen rays was studied with the help of a number of stereoscopic pictures of tracks in a Wilson cloud chamber by F. W. Bubb* and F. Kirchner,† who used a considerably greater number of individual measurements. Even the first experiments, those of F. W. Bubb, showed that the direction of the electric vector exhibits a clearly marked favoured direction for the photo-electric emission—in qualitative agreement with the general lines of reasoning of p. 186. Quantitatively, the measurements of Kirchner gave a direction-distribution which agreed, within the limits of error of the statistical method, with the $\cos^2 \theta$ law of eqn. (13). A further agreement with eqn. (19) occurs in the experimental result that the form of the “azimuthal” distribution curve (angle ϕ) is independent of the wave-length in the region between $\lambda = 0.3$ and 0.8 \AA. , and from the ionisation potential in the region between $J = 300$ to 3000 volts. As to the “longitudinal” distribution curve (angle θ) the $\sin^2 \theta$ law for J_θ , which was first deduced in eqn. (13b) for polarised rays, but which obviously applies equally well to unpolarised rays, was experimentally established by P. Auger‡ for unpolarised rays.

The fact that the centre of gravity of the photo-electric emission moves further and further “forwards” as the wave-length decreases was early the object of a whole series of electrometric measurements. The advance of the maximum was established quantitatively only quite recently by those methods which allow elementary processes to be registered individually, namely by means of the Geiger point-counter § (*Spitzenzähler*), and Wilson’s cloud-chamber.||

By eqn. (21) or (27) the bisecting cone should be displaced forwards by an amount proportional to the velocity v with which the electron leaves the atom. This is in qualitative agreement with the experimental result that the maximum moves further and further forward according as the wave-length of the incident radiation becomes smaller, and also that it moves less forward the greater the work of ionisation to be performed on the atom (Bothe, Auger).

Eqn. (21) also gives the correct order of magnitude *quantitatively* for the position of the bisecting cone, as is seen from Fig. 20. This represents the forward angle $\pi/2 - \theta_0$ (*Voreilwinkel*) as a function of the velocity with which the electrons have in each case left the atom; the continuous straight line corresponds to the theoretical result given by eqn. (21). The earlier measurements almost all lie below this straight line. In the figure only the most recent measurements have been inserted; these were

* F. W. Bubb, *Phys. Rev.*, **23**, 137 (1924).

† F. Kirchner, *Ann. d. Phys.*, **83**, 521 (1927).

‡ P. Auger, *Journ. d. Phys.*, **8**, 85-112 (1927).

§ W. Bothe, *Zeitschr. f. Phys.*, **26**, 59 (1924).

|| Williams, Nuttal, and Barlow, *Proc. Roy. Soc.*, **121**, 611 (1928); P. Auger, *Comptes Rendus*, **186**, 758 (1928); **187**, 141 (1928); **188**, 447 (1929).

carried out with homogeneous rays reflected from crystals. They come notably near the theoretical* continuous line in the figure.

We should therefore like to assume that our numerical factor 2 by which the wave-mechanical calculation deviates from the interpretation in terms of radiation pressure has a basis in reality. The dotted straight line corresponds to the latter interpretation, which seems to suggest itself more readily physically.

An agreement with the theoretical assertion contained in eqn. (23) is also to be seen in the experimental result, particularly emphasised by Auger, that the form of the distribution is independent of the position of the maximum and of the amount of the work of ionisation. The theory does not, however, comprehend some subsidiary maxima that have been found by Auger and Bothe in the longitudinal curve of distribution; these maxima do not yet seem to be experimentally secure.

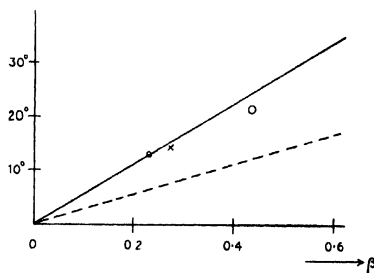


FIG. 20.—Forward displacement of the bisecting angle (*Halbwinkel*) of the photo-electric emission.

Continuous line: calculated by wave-mechanics.

Broken line: expected results according to the elementary view.

° Auger, × Williams, Nuttal and Barlow.

Our calculation in C was only a first approximation, since we broke off the expansion of the exponential function (16a) at the first power of r/λ . In the case of hard Röntgen rays this is, strictly, no longer admissible, as then the wave-length λ becomes of the same order of magnitude as the radius of the K-shell of the atom in question. In the case of γ -rays the power-expansion becomes illusory altogether; we are then driven to integrate the exponential function exactly.

§ 5. Diffraction at a Single Obstacle. Collision Problems of two Particles

To link up closely with the preceding sections it would seem most natural to continue with perturbation problems and to treat the Compton

* We wish to call attention to the fact that in the German edition of this book the factor 9/5 occurred by an oversight in place of the correct factor 2 [cf. above in eqns. (27) and (28)]. Corresponding to this the theoretical straight line in Fig. 20 of the German edition lay a little lower than in our present figure.

effect by the method of wave-mechanics. We find it preferable, however, to postpone the treatment of further perturbation problems and first to occupy ourselves with the simpler diffraction problems. Incidentally we shall pave the way mathematically for the later treatment of the Compton effect.

In the diffraction problems of optics we distinguish between *diffraction by a single object* (slit, aperture, disc) and *diffraction by a system of regularly arranged objects* (grating lines, crossed gratings, space lattices). In the present section we shall deal with diffraction due to a single object, which we shall imagine to be an atom, nucleus or molecule; in A the diffracted wave will first signify a stream of α -particles, in B and C a stream of electrons. In the succeeding section we shall then pass on to the diffraction of electron waves by a regularly arranged system of diffracting objects.

A. Deflection of α -rays by Atoms and Atomic Nuclei

Micromechanics claims that the wave-equation is the key to phenomena which are usually treated in a less refined way by the methods of ordinary mechanics. A characteristic example* is given by the experiments referred to in the title of this sub-section, which in the hands of Rutherford first led to the idea of the nuclear model and so form the foundation of the whole of atomic physics.

We imagine an α -particle (or, in the sense of § 8, Chap. I, rather a stream of α -particles) incident from the direction of the negative x -axis. By eqns. (14) and (16), § 1, Chap. I, this phenomenon is described micromechanically by the wave-function

$$\psi_0 = e^{ikx}, \quad k = \frac{2\pi}{\lambda}, \quad \lambda = \frac{h}{Mv} = \frac{h}{\sqrt{2ME}} \quad . \quad . \quad (1)$$

M is the mass of the α -particle, $E = Mv^2/2$ its energy and at the same time the proper value of our problem. As we here wish to consider only ordinary scattering without the transfer of energy to the atom, that is, only *elastic* collisions, E does not become changed by the disturbance emanating from the atom. Thus we simplify our problem by regarding its proper value E as known from the outset.

Let the disturbing cause be a neutral atom with a Z -fold charged nucleus at the zero point (origin of the co-ordinate system); we imagine its Z electrons—in contravention of all the rules of the periodic system—as united in the K -shell and without influence on one another. The effect of these electrons at external points is then represented wave-mechanically by the Z -fold electrop cloud of hydrogen in the ground state, that is, by its

* G. Wentzel, *Zeitschr. f. Phys.*, 40, 590 (1927). In relationship with more general problems: M. Born, *Göttinger Nachr.*, 1926, p. 146, and W. Elsasser, *Zeitschr. f. Phys.*, 45, 522 (1927).

which allows us to neglect the product $V\psi_1$. Eqn. (5) then becomes [cf. the explanation of k in eqn. (1)]

$$\Delta\psi_0 + k^2\psi_0 + \Delta\psi_1 + k^2\psi_1 - k^2\frac{V}{E}\psi_0 = 0.$$

Since here the first two terms, taken by themselves, vanish, we get for ψ_1 the differential equation,*

$$\Delta\psi_1 + k^2\psi_1 = k^2\frac{V}{E}\psi_0 \quad . \quad . \quad . \quad (6)$$

which has become non-homogeneous, owing to the addition of the perturbation function V .

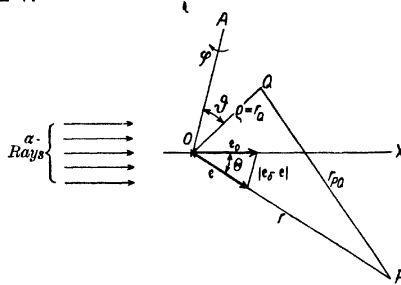


FIG. 21.—Scattering of a beam of α -rays by an atom at O.

The method of integrating this equation is shown by the following comparison:

Potential Theory

$$\Delta\phi = -4\pi\rho,$$

$$\phi = \int \frac{\rho}{r} d\tau.$$

Vibration Theory

$$\Delta\psi + k^2\psi = -4\pi\sigma,$$

$$\psi = \int \frac{\sigma}{r} e^{ikr} d\tau.$$

The integral expression for ϕ is the well-known form corresponding to Newton's law. In the same way, namely from Green's theorem, eqn. (14) of § 5, Chap. I, we deduce the integral form of ψ , by using as the characteristic function, instead of $1/r$ (the solution of $\Delta\phi = 0$), e^{ikr}/r (the solution of the corresponding homogeneous equation $\Delta\psi + k^2\psi = 0$). Using the characteristic function in (6) we get

$$\psi_P = -\frac{2\pi e^2 Z}{\lambda^2 E} \int \left(\frac{1}{r_Q} + \frac{Z}{a} \right) e^{-\frac{2Z}{a} r_Q + ikx_Q} \frac{e^{ikr_{PQ}}}{r_{PQ}} d\tau \quad . \quad . \quad (7)$$

P (cf. Fig. 21) is the reference point (*Aufpunkt*) for which ψ_1 is to be calculated, Q is the point of integration. For in the expression of

$$\sigma = -\frac{k^2}{4\pi} \frac{V}{E} \psi_0$$

* In the same way we get for the perturbation ψ_n of the n th order the recurrence formula

$$\Delta\psi_n + k^2\psi_n = k^2 \frac{V}{E} \psi_{n-1}.$$

we have to replace r by r_Q = distance OQ and α by x_Q . On the other hand, r denotes in the expression of the characteristic function e^{ikr}/r the distance PQ, which we have denoted by r_{PQ} . Finally, let r , without an index, be the distance OP. We are interested only in the positions of P for very great values of r . We then have, in vectorial language,

$$\left. \begin{aligned} r_{PQ} &= |\mathbf{r} - \mathbf{r}_Q| = [\mathbf{r}^2 - 2(\mathbf{r}\mathbf{r}_Q) + \mathbf{r}_Q^2]^{1/2} \\ &= r \left(1 - \frac{1}{r}(\mathbf{r}\mathbf{r}_Q) + \dots \right) = r - (\mathbf{e}, \mathbf{r}_Q) \end{aligned} \right\} \quad (8)$$

where \mathbf{e} denotes unit vector in the direction OP; further we introduce the unit vector \mathbf{e}_0 in the direction OX (cf. Fig. 21), by which x_Q is expressed as follows:

$$x_Q = (\mathbf{e}_0, \mathbf{r}_Q) \quad (8a)$$

It is clear that the 0th approximation $r_{PQ} = r$ will suffice in the denominator of (7), and we need use the first approximation (8) only in the exponent of e . Eqn. (7) then becomes

$$\psi_P = - \frac{2\pi e^2 Z}{\lambda^2 E} \frac{e^{ikr}}{r} \int \left(\frac{1}{r_Q} + \frac{Z}{a} \right) e^{-\frac{2Z}{a} r_Q + ik(\mathbf{e}_0 - \mathbf{e}, \mathbf{r}_Q)} d\tau. \quad (9)$$

The integration may now easily be performed in an appropriately chosen system of polar co-ordinates, whose polar axis OA lies parallel to the difference vector $\mathbf{e}_0 - \mathbf{e}$. Let the co-ordinates of the point of integration Q in this system be

$$\rho = r_Q, \quad \theta = \angle AOQ, \quad \phi = \angle \text{about OA}.$$

Then we get simply

$$(\mathbf{e}_0 - \mathbf{e}, \mathbf{r}_Q) = |\mathbf{e}_0 - \mathbf{e}| \rho \cos \theta.$$

According to the figure, however, $|\mathbf{e}_0 - \mathbf{e}|$ is the base of an isosceles triangle, whose angle at O is equal to Θ ($\Theta = \angle XOP$ = angle of diffraction or scattering angle) and both of whose sides = 1. Thus,

$$|\mathbf{e}_0 - \mathbf{e}| = 2 \sin \frac{\Theta}{2}, \quad (\mathbf{e}_0 - \mathbf{e}, \mathbf{r}_Q) = 2\rho \sin \frac{\Theta}{2} \cos \theta. \quad (9a)$$

By performing the integration with respect to ϕ immediately and writing ψ_1 again in place of ψ_P , (9) becomes

$$\psi_1 = C_I \cdot I + C_{II} \cdot II \quad (10)$$

$$C_I = - \frac{4\pi^2 e^2 Z}{\lambda^2 E} \frac{e^{ikr}}{r}, \quad C_{II} = - \frac{4\pi^2 e^2 Z^2}{\lambda^2 E a} \frac{e^{ikr}}{r},$$

$$I = \int_0^\pi \sin \theta d\theta \int_0^\infty \rho d\rho e^{-(c - id \cos \theta)\rho},$$

$$II = \int_0^\pi \sin \theta d\theta \int_0^\infty \rho^2 d\rho e^{-(c - id \cos \theta)\rho}.$$

$$c = \frac{2Z}{a}, \quad d = 2k \sin \frac{\Theta}{2} \quad (10a)$$

Integration with respect to ρ gives for I and II, respectively,

$$\frac{1}{(c - id \cos \theta)^3} \quad \frac{2}{(c - id \cos \theta)^3},$$

and integration with respect to θ , if we make $y = c - id \cos \theta$,

$$I = \frac{1}{id} \int_{c-id}^{c+id} \frac{dy}{y^3} = \frac{1}{id} \left(\frac{1}{c-id} - \frac{1}{c+id} \right) = \frac{2}{c^2 + d^2},$$

$$II = \frac{2}{id} \int_{c-id}^{c+id} \frac{dy}{y^3} = \frac{2}{2id} \left(\frac{1}{(c-id)^2} - \frac{1}{(c+id)^2} \right) = \frac{4c}{(c^2 + d^2)^2}.$$

We first consider only the first component C_1 of ψ_1 ; on account of (10a) we have

$$|\psi_1| = \frac{8\pi^2 e^2 Z}{\lambda^2 E r} \frac{1}{4k^2 \sin^2 \frac{\Theta}{2} + \frac{4Z^2}{a^2}}.$$

By our definition of ψ_0 in (1), we have $|\psi_0| = 1$, hence, by taking out the factor $4k^2$ in the denominator of the last equation and observing that $k\lambda = 2\pi$, we may write

$$\left| \frac{\psi_1}{\psi_0} \right|^2 = \left(\frac{e^2 Z}{2Er} \right)^2 \frac{1}{\left(\sin^2 \frac{\Theta}{2} + \alpha^2 \right)^2} \quad \dots \quad (11)$$

with the abbreviation

$$\alpha = \frac{\lambda}{a} \frac{Z}{2\pi} \quad \dots \quad (11a)$$

Eqn. (11) may then be interpreted as the ratio of the number of particles $|\psi_1|^2$ deflected through an angle Θ , to the number of incident particles $|\psi_0|^2$ and agrees exactly with the formula of classical mechanics (see below), according to which Rutherford originally worked out the results of his experiments—on the assumption that we may neglect the correction term α^2 .

We now test the validity of this last assumption. By (1) we have

$$\lambda = \frac{h}{Mv} = \frac{h}{\mu v} \frac{\mu}{M} = \frac{3.6 \cdot 10^{-9}}{7000} \quad \dots \quad (11b)$$

where the denominator 7000 denotes the ratio M/μ (mass of the helium atom divided by the mass of the electron), and the numerator denotes the quantity $h/\mu v$ with $v = 2 \cdot 10^9$. Hence, for $a = 0.53 \cdot 10^{-8}$ (cf. I, p. 212)

$$\frac{\lambda}{a} \sim 10^{-4} \quad \dots \quad (11c)$$

and for $Z = 79$ (Au)

$$\alpha \sim 1.2 \cdot 10^{-2}.$$

Thus we should have to consider extremely small angles Θ if the correction term α^2 is to be taken into consideration with $\sin^2 \frac{\Theta}{2}$; for all values of Θ accessible to observation it may be neglected.

For the same reason the second term in (10) is to be neglected in

comparison with the first. According to the data just given we actually see that

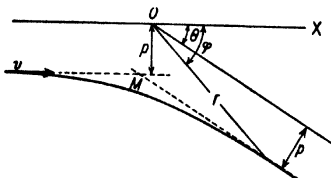
$$\frac{C_{II}}{C_I} \frac{II}{I} = \frac{Z}{a} \frac{2c}{c^2 + d^2} \sim \frac{4Z^2}{a^3} \frac{1}{4k^2 \sin^2 \frac{\Theta}{2}} = \frac{\alpha^2}{\sin^2 \frac{\Theta}{2}} \ll 1$$

(if Θ is not very small). Hence we may compress our final result practically under the conditions of α -ray scattering into the formula

$$\left| \frac{\psi_1}{\psi_0} \right|^2 = \left(\frac{e^2 Z}{2Er} \right)^2 \frac{1}{\sin^4 \frac{\Theta}{2}} \quad (12)$$

It now only remains to derive this formula from purely corpuscular theory. The following method which links up with I, Chap. VII, p. 472, seems particularly convenient. Let the continuation of the original direction of the particle pass at a distance p from the origin, that is, the position of the deflecting nucleus.

We write down the law of areas and the law of conservation of energy in plane polar co-ordinates (cf. Fig. 22):



$$r^2 \dot{\phi} = pv, \\ \frac{M}{2} (\dot{r}^2 + r^2 \dot{\phi}^2) + \frac{2e^2 Z}{r} = E.$$

Fig. 22.—Deflection of an α -particle by a nucleus according to classical mechanics.

variable. This gives

$$\frac{M}{2} p^2 v^2 \left[\left(\frac{ds}{d\phi} \right)^2 + s^2 \right] + 2e^2 Z s = E.$$

To arrive at a linear differential equation, differentiate with respect to ϕ and cancel the common factor $ds/d\phi$. We get

$$\frac{d^2 s}{d\phi^2} + s = -C, \quad C = \frac{2e^2 Z}{M p^2 v^2} = \frac{e^2 Z}{p^2 E}.$$

The general solution of this equation is

$$s = -C + A \cos \phi + B \sin \phi \quad (13)$$

The initial state $\phi = \pi, s = 0$ (cf. the figure) makes

$$A = -C \quad (14)$$

From this, if we write in place of (13)

$$1 = -C(r+x) + By \quad (13a)$$

we get likewise for the initial asymptote

$$y = p, \quad r + x = r - |x| = 0,$$

and so by (13a)

$$B = \frac{1}{p} \quad (15)$$

Let the final asymptote be inclined at the angle $\phi = \Theta$. Then (13) shows, when combined with (14) and (15),

$$\begin{aligned} C(1 + \cos \Theta) &= \frac{1}{p} \sin \Theta \\ \cot \frac{\Theta}{2} &= \frac{1}{Cp} = \frac{pE}{e^2Z} \end{aligned} \quad (16)$$

(13) is the polar equation of the hyperbola, O its (external) focus; the centre M has been shown in the figure as the intersection of the asymptotes. Let the number of particles that fall on unit area at right angles to the direction of the primary rays be n , so that $n\pi p^2 = N$, say, fall on the surface πp^2 . By (16) we may therefore write

$$N = n\pi \left(\frac{e^2Z}{E} \right)^2 \cot^2 \frac{\Theta}{2}.$$

Hence the number of particles that are deflected through an angle between Θ and $\Theta + d\Theta$ is

$$dN = n\pi \left(\frac{e^2Z}{E} \right)^2 \frac{\cot \frac{\Theta}{2} d\Theta}{\sin^2 \frac{\Theta}{2}}.$$

These particles fall at a distance r on to the zone of a sphere of area

$$dF = 2\pi r^2 \sin \Theta d\Theta.$$

Hence the number that fall on unit area at this distance is

$$\frac{dN}{dF} = \frac{n(e^2Z)^2}{2(rE)^2} \frac{\cot \frac{\Theta}{2}}{\sin \Theta \sin^2 \frac{\Theta}{2}}.$$

If we here replace $\sin \Theta$ by $2 \sin \frac{\Theta}{2} \cos \frac{\Theta}{2}$ and form the ratio of the deflected to the incident particles, both being calculated for unit area, that is if we calculate

$$\frac{1}{n} \frac{dN}{dF}$$

we get exactly eqn. (12).

B. Comparison with Experiment. Application of the Preceding Results to Electron Rays

The first question that is to be decided is whether the deviation, expressed by α^2 , of our formula (11) from Rutherford's formula (12) is of *wave-mechanical* or *statistical* origin, that is, whether it is due to the refined treatment of mechanics or to the statistical interpretation of the atomic model which Schrödinger's formula (2) entails. The answer to this question is that the α^2 -term in question is of *statistical* origin. This follows from an investigation of W. Gordon,* which deals with the

* *Zeitschr. f. Phys.*, **43**, 180 (1928); cf. also J. R. Oppenheimer, *ibid.*, **43**, 418 (1927).

collision of an unaccompanied nucleus (without its neutralising electronic envelope) with a stream of parallel corpuscles and leads rigorously to Rutherford's formula without the α^2 -term. Thus in the case of the collision of two particles that act on each other purely according to Coulomb's law, wave-mechanics agrees exactly with classical mechanics.*

Concerning the methodological features involved we must make the following remarks. In the diffraction of light two supplementary methods have long been known: on the one hand, the method of Huygen's principle, on the other, that of expansion in series, the former being suitable for small wave-lengths, the latter for long wave-lengths.

Gordon's method is expressly based on the method of expansion in series and is worked out exactly; in contrast with this, our procedure, which follows along the lines of Wentzel, corresponds in a certain sense to Huygen's principle and is in the first place an approximation. The exact working out of this method by means of approximations of the second and higher order has been given by Born and Elsasser, cf. the reference at the beginning of section A.

We now turn to experiments with beams of electrons. It is clear that the correction term α^2 here plays a much more important part than in experiments with α -rays. For in our estimate of λ in eqn. (11b) the factor 7000 now drops out of the denominator, so that λ/a , eqn. (11c), becomes of the order of magnitude 1. For α there also results according to the particular value of the nuclear charge Z and the velocity v the order of magnitude 1.

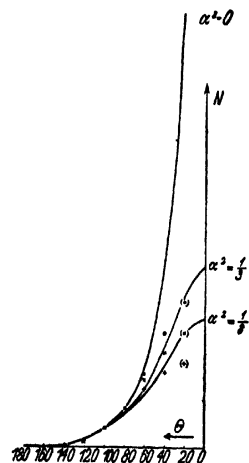


FIG. 23.—Deflection of α -rays in argon.

- Velocity > 15 kV.
- + Velocity < 10 kV.
- × Arithmetic mean of both.

Mr. F. Kirohner has kindly looked through the statistical data published by him earlier† to see whether the α^2 -term makes itself felt in them. The cathode rays involved were of velocities ranging from 10 to 40 kilovolts, in particular in argon, and the results were obtained by Wilson's cloud-chamber method. In counting the collisions only strong deflections, 20° and more, were considered; these are deflections in which the electron approaches sufficiently near the nucleus, or, as we may now say, penetrates into the K-shell. The

* Disagreements which have been found for fast α -particles in the Cavendish Laboratory appear to be due rather to some cause in the nuclear structure than in a failure of classical mechanics.

† Ann. d. Phys., 83, 939 (1927).

deciding velocity of the electron in an individual collision was estimated from the length of path of the emergent electron in the Wilson photograph.

In Fig. 23 the abscissæ denote the angles of deflection Θ between 0 and 180° (strong deflections on the left); the ordinates are the numbers N of all particles that are deflected by more than Θ . The formula for this integral number of particles is, by (10) and the associated calculations,

$$N = C \int_{\Theta}^{\pi} \left(\frac{1}{\sin^2 \frac{\Theta}{2} + \alpha^2} + \frac{\alpha^2}{\left(\sin^2 \frac{\Theta}{2} + \alpha^2 \right)^2} \right)^2 \sin \Theta \, d\Theta.$$

The integration may easily be performed; in particular, we get for $\alpha^2 = 0$ the expression $N = 2C \cot^2 \frac{\Theta}{2}$. The corresponding curve, which is thus in agreement with Rutherford's law, has an infinite maximum for $\Theta = 0$. As we see, the observations lie far below this curve. The averaged observations (\times) may be regarded as following the theoretical curve for $\alpha^2 = \frac{1}{4}$; the observations at higher velocities (\circ) lie above this, those at lower velocities ($+$) lie below this, as is to be expected. The theoretical value of α^2 comes out according to eqn. (11a), for argon ($Z = 18$) and for a velocity of 15 kV to about a third. According to what we said at the beginning of this section, this deviation from the curve for $\alpha^2 = 0$ is not a wave-mechanical diffraction effect but a statistical screening effect of the K-electrons, which are near the nucleus. This screening effect could also be calculated according to the older theory by averaging suitably over all possible positions of the K-orbits. But it is immediately obvious how much simpler and more direct this calculation becomes in Schrödinger's theory which has taken up the statistics of all possible orbits in its foundations and hence furnishes directly the necessary means for discussing the observations.

C. Collisions between Electrons and Gas Molecules

E. Fermi* gives a particularly instructive and simple example of the application of wave-mechanics to collision phenomena. To allow it to be pictured in three dimensions, he makes the following idealisation: electrons are incident from any particular direction in the xy -plane as a parallel beam; they encounter gas molecules which can rotate only about an axis fixed in space, say perpendicular to the xy -plane. Thus we have three degrees of freedom: the co-ordinates xy of the electron and the angle of rotation ϕ of the molecule. We introduce the latter as a third

* Zeitschr. f. Phys., 40, 899 (1927).

rectilinear co-ordinate

$$z = \sqrt{\frac{J}{\mu}} \phi \quad (17)$$

where J is the moment of inertia of the molecule about its axis of rotation, μ the mass of the electron. The wave-equation of this problem is, then, just like the wave-equation for a point-mass in the three-dimensional space of the x , y , and z :

$$\Delta\psi + \frac{8\pi^2\mu}{h^2}(E - V)\psi = 0 \quad (18)$$

V is the potential energy of the mutual action between the electron and the molecule, E the total energy of both together.

We first see that, by (17), z has the dimensions of a length, since the dimensions of J is mass multiplied by the square of a length. We next remark that the kinetic energy of the electron + the molecule is

$$\frac{\mu}{2}(\dot{x}^2 + \dot{y}^2) + \frac{J}{2}\dot{\phi}^2 = \frac{\mu}{2}(\dot{x}^2 + \dot{y}^2 + \dot{z}^2).$$

So our definition of z leads to the same kinetic energy and hence also (cf. § 11 A, Chap. I) the same wave-equation as for the individual point-mass in three dimensions.

Let the molecule be situated at the point $x = y = 0$; according to its phase of rotation (*Umdrehungsphase*), however, it will be represented by a varying point on the z -axis. When the electron is at a great distance from the molecule $V = 0$ and $E = E_1 + E_2$, where E_1 denotes the kinetic energy of the electron, E_2 that of the molecule. The coefficient of ψ in the wave-equation, which corresponds to the square of the refractive index (cf. eqn. (8), § 1, Chap. I), is then constant. Thus we have an optically homogeneous and isotropic medium, *except in the vicinity of the z -axis*. Here V is periodically variable, simultaneously with the phase of the motion of the molecule. The period is [cf. (17)]

$$a = 2\pi\sqrt{\frac{J}{\mu}} \quad (19)$$

In the theory of waves this denotes a *linear grating*. Actually, a periodic change in the index of refraction, such as has just been described, is exactly the same in wave-mechanics as a linear-succession of diffracting apertures or of optical obstacles of some kind or other. So we apply the ordinary grating formula, eqn. (1), § 1, Chap. IV,

$$\alpha - \alpha_0 = n \frac{\lambda}{a} \quad (20)$$

Here λ is the wave-length of the de Broglie wave corresponding to our moving system: we get its value most simply from the wave-equation. n is the order of the diffraction phenomenon; α_0 and α denote the cosines of the angles of incidence and diffraction, both measured from the

z -direction. We do not mean, however, the angle of incidence of the arriving electron with respect to the z -axis, which by our initial assumption is equal to 90° , like the angle at which the diffracted electron leaves, but the corresponding angles for the de Broglie wave corresponding to our total system. We also best obtain these angles from the wave-equation or its integral.

At a great distance from the z -axis (18) becomes, on account of $V = 0$,

$$\Delta\psi + k^2\psi = 0, \quad k = \frac{8\pi^2\mu E}{h^2} \quad (21)$$

This determines λ ; for

$$\lambda = \frac{2\pi}{k} = \frac{h}{\sqrt{2\mu E}} \quad (22)$$

We integrate (21) as in the optical problem of the plane light-wave and write

$$\psi = A e^{ik(x \sin \theta \cos \chi + y \sin \theta \sin \chi + z \cos \theta)} \quad (23)$$

θ is the angle the wave-normal makes with the z -axis, so that $\cos \theta = \alpha_0$ or $= \alpha$, according to whether we consider the incident or the diffracted wave. χ determines the direction of the incident or the diffracted electron, respectively, in the xy -plane.

But we may also separate eqn. (21) by factorising ψ as $\psi_1 \psi_2$ and writing $E_1 + E_2$ for E , and then making ψ_1, ψ_2 obey the equations separately:

$$\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) \psi_1 + k_1^2 \psi_1 = 0, \quad k_1^2 = \frac{8\pi^2\mu}{h^2} E_1 \quad (24a)$$

$$\frac{\partial^2}{\partial z^2} \psi_2 + k_2^2 \psi_2 = 0, \quad k_2^2 = \frac{8\pi^2\mu}{h^2} E_2 \quad (24b)$$

Our justification for this is that E_1 belongs to the electronic motion, E_2 to the molecular rotation, and that at great distances these components are independent of each other when the electron and molecule are at a great distance from each other.

Integration of (24a) and (24b) gives

$$\left. \begin{aligned} \psi_1 &= A_1 e^{ik_1(x \cos \chi + y \sin \chi)} \\ \psi_2 &= A_2 e^{ik_2 z} \end{aligned} \right\} \quad (25)$$

where χ has the same meaning as before. A comparison of the coefficients of x, y, z in (23) and (25) gives

$$k \sin \theta = k_1, \quad k \cos \theta = k_2 \quad (26)$$

On account of (21) and (24a, b) this is equivalent to

$$\sin \theta = \sqrt{\frac{E_1}{E}}, \quad \cos \theta = \sqrt{\frac{E_2}{E}} \quad (27)$$

The periodicity of V leads to the wave-state also being periodic in the

204 Chapter II. Perturbations. Diffraction Problems

z -direction, with the period a , eqn. (19). So we shall demand for the incident and the diffracted wave in turn

$$k_2 = \frac{2\pi}{a}m = \sqrt{\frac{\mu}{J}}m, \quad k_2^* = \frac{2\pi}{a}m^* = \sqrt{\frac{\mu}{J}}m^*,$$

where m and m^* are whole numbers. From this we get for the rotational energy E_2 , E_2^* respectively before and after the collision by eqn. (24b),

$$E_2 = \frac{m^2 h^2}{8\pi^2 J}, \quad E_2^* = \frac{m^{*2} h^2}{8\pi^2 J} \quad . \quad . \quad . \quad (28)$$

This is the Deslandres term [I, Chap. VII, p. 415], and it appears with integral, not half-integral rotation-quantum numbers m , m^* corresponding to the fact that we are considering a rotator with a *fixed* axis [§ 3, C, Chap. I].

The value of $\cos \theta$ before and after the collision now follows from (27) and (28). If, as above, we call it α_0 and α we have

$$\alpha_0 = \frac{mh}{\sqrt{8\pi^2 J E}}, \quad \alpha = \frac{m^* h}{\sqrt{8\pi^2 J E}} \quad . \quad . \quad . \quad (29)$$

If we finally insert these values in the diffraction condition (20) and also the values (19) and (22) for α and λ , it follows that

$$m^* - m = n \quad . \quad . \quad . \quad (30)$$

Thus the collision involves quanta. The rotator gains or loses $|n|$ whole rotation quanta, according as n is positive or negative.

At the same time, on account of the constancy of the total energy E the electron loses or gains the amount of energy determined by (28):

$$E_1^* - E_1 = - (E_2^* - E_2).$$

We have a collision of the first or second kind (cf. Atombau, Chap. VII, p. 530) according as this amount is negative or positive, that is, according as the molecule is excited by the electron or the electron by the molecule in the collision. In the case $m = m^$, $E_1 = E_1^*$ the collision is elastic.*

The angles χ and χ^* of the electronic motion of course remain indeterminate; for in our linear lattice all azimuths within the diffracting cone described about the z -axis (apertures being θ_1^* , θ_2^* , . . . θ_n^* for the first, second . . . n th order) are equivalent.

The particular elegance of this example taken from Fermi consists in the circumstance that—just as in the theory of the diffraction grating—no particular assumption about the nature of the grating (here, about the potential energy V) is necessary, so long as no quantitative data concerning the relative intensity of the different diffraction patterns are required (here, the number of the different elastic and inelastic collisions that occur).

§ 6. Diffraction of Matter by Gratings. Interference of Electron-waves in Crystals

We have already used the theory of gratings in the preceding section to treat the diffraction of electronic rays by a rotating molecule, by the method of wave-mechanics. In this case the diffracting object was only a single molecule, not a regularly arranged set of diffracting objects. We now turn to the proper theory of gratings, that is, to the *diffraction of matter-waves in crystals*.

The first experiments on this fascinating subject fall to the credit of Davisson and Germer. The investigation concerned the diffraction of electrons from a single crystal (*Einkristall*) of nickel. The observed diffraction phenomena are true Laue points (*Flecke*), the arrangement of which faithfully reproduces the symmetry of the single crystal as in the diffraction of Röntgen rays. The later experiments on thin metal foils of micro-crystalline structure are equally impressive; the earliest experiments of this type are due to G. P. Thomson. The observed diffraction phenomena correspond exactly to the Debye-Scherrer-Hull rings of Röntgen ray analysis. We may here mention that the electrical conductivity of metals also gives information about the diffraction and interference of the conduction electrons in the lattice of metallic ions seeing that the newer statistics of electrons according to wave-mechanics allows the temperature-dependence of the conductivity to be calculated * satisfactorily from the same formulæ as are used for the scattering of Röntgen rays in the crystal lattice. But as the latter problem lies outside our scope, we shall restrict ourselves to the theory of the first-mentioned experiments.

A. The Experiments of Davisson and Germer †

A beam of electrons of uniform velocity in a vacuum falls perpendicularly on to the surface of the nickel crystal (octahedral surface 111); observations were made on the side of incidence by an electrometric method. The potential of the electron tube varied in the different experiments from 30 to 370 volts.

At 100 volts the velocity is, as we know,

$$v = \sqrt{\frac{2e}{m}V} \cdot 10^8 = 6 \cdot 10^8 \text{ cm./sec.}$$

By de Broglie's formula

$$\lambda = \frac{h}{mv}$$

* W. V. Houston, *Zeitschr. f. Phys.*, **43**, 449 (1928); cf. also A. Sommerfeld, *Naturwiss.*, **13**, 374 (1928).

† *Phys. Rev.*, **30**, 705 (1927); cf. also *Nature*, **119**, 558 (1927), and *Proc. Nat. Acad.*, **14**, 817 (1928), as well as the earlier experiments of Davisson and Kunsman, *ibid.*, **22**, 242 (1928). The theory given in the text is due to H. Bethe; cf. the provisional notions in the *Naturwiss.*, **15**, 786 (1927) and, in particular **16**, 888 (1928). As early as 1925, W. Elsasser [*Naturwiss.*, **13**, 711 (1925)] interpreted the experiments of Davisson and Kunsman by means of the interference of de Broglie waves.

the axis perpendicular to the bounding octahedral surface, which we shall call θ , ϕ and θ_0 , ϕ_0 , respectively, inside the crystal, and Θ , Φ and Θ_0 , Φ_0 respectively outside,* we have

$$\left. \begin{aligned} \alpha &= \sin \theta \cos \phi & \beta &= \sin \theta \sin \phi & \gamma &= \cos \theta \\ \alpha_0 &= \sin \theta_0 \cos \phi_0 & \beta_0 &= \sin \theta_0 \sin \phi_0 & \gamma_0 &= \cos \theta_0 \end{aligned} \right\} \quad (4)$$

According to the law of refraction, we have for the passage from air to crystal or crystal to air respectively,

$$\frac{\sin \Theta_0}{\sin \theta_0} = \frac{\sin \Theta}{\sin \theta} = \mu, \quad \Theta_0 = \phi_0, \quad \Phi = \phi. \quad (5)$$

If we substitute this in (4) and (3) and take (2) into account, we get

$$\left. \begin{aligned} a \left(\frac{\sin \Theta}{\mu} \cos \Phi - \frac{\sin \Theta_0}{\mu} \cos \Phi_0 \right) &= n_1 \frac{\Lambda}{\mu} \\ b \left(\frac{\sin \Theta}{\mu} \sin \Phi - \frac{\sin \Theta_0}{\mu} \sin \Phi_0 \right) &= n_2 \frac{\Lambda}{\mu} \\ c \left(\sqrt{1 - \frac{\sin^2 \Theta}{\mu^2}} - \sqrt{1 - \frac{\sin^2 \Theta_0}{\mu^2}} \right) &= n_3 \frac{\Lambda}{\mu} \end{aligned} \right\} \quad (6)$$

In the first two equations μ may be divided out; they have then exactly the same form in Λ , Θ , Φ as eqns. (3) in λ , θ , ϕ . But this means that the diffraction at the crossed grating which is formed in the bounding surface by the directions 1 and 2, takes place in exactly the same way as for Röntgen rays of wave-length Λ , that is, as if the index of refraction were $\mu = 1$. In the third eqn. (6), on the other hand, μ occurs in an essential way; it differs from the third eqn. (3). From this it follows that the crystal dimensions appear distorted in the direction of depth, 3, if we calculate them as for Röntgen rays, that is without taking into account the index of refraction μ . This is just what Davisson and Germer deduced from their experiments: *agreement* in the general relationships between wave-length and inclination, as follows in the case of surface lattices from the first two Laue equations, but *difference* in the choice of wave-lengths, which are determined only by the help of the third Laue equation, that is by the space-lattice effect. This last disagreement with the corresponding experiment in Röntgen rays has provisionally been formulated by Davisson and Germer as being equivalent for slow electrons to an apparent contraction of the crystal lattice in the direction of the depth by the amount 70 per cent.

We shall next get an idea of the true reason of this apparent distortion, that is, of the magnitude of the index of refraction μ . Schrödinger's equation for the electron in uniform motion serves this purpose; we shall write it in the form:

$$\Delta \psi + \frac{8\pi^2 me}{h^2} (E + V) \psi = 0 \quad (7)$$

* According to what was said above about the direction of incidence $\Theta = 0$ in the experiments of Davisson and Germer. But it keeps things clearer to calculate in the sequel with an arbitrary Θ_0 .

The change in the manner of writing it now as compared with formerly consists in the addition of the factor e to the left-hand side and reversal of the sign of V . This means that eE now denotes the total energy of the electron and $-eV$ its potential energy at the point in question. If we measure the energy of the electron, as usual, in volts, then E is directly the number of volts and V the electric potential likewise measured in volts (except for the factor 10^8 or $\frac{1}{300}$ according as e is measured in e.s.u. or e.m.u.; this factor is of no account in the sequel). E , being the proper-value parameter of the whole problem, has the same value in *vacuo* as in the crystal. V , on the other hand, suffers a sudden change at the surface of separation, or, as we shall assume simply in the sequel a discontinuous change. In a vacuum (outside the crystal) V must be set equal to 0; in the lattice of the crystal V will be periodically variable, but so that the mean value V_0 of V is *positive*. For the electron is in stable equilibrium in the metal. Hence its potential energy here, compared with the outside of the crystal, must be *negative* which according to the above method of introducing V is equivalent to saying that V_0 becomes *positive*. We can consider the interior of the metal as an energy-pocket (*Mulde*) for the negative electron. Work must be done to lift the electron out of the pocket (Richardson effect). Conversely, the velocity of the electron *increases* if it is fired from outside into the metal. Since, by de Broglie's fundamental equation, the velocities are in the inverse ratio of the wave-lengths, the wave-length in the interior of the metal becomes *smaller* than the wave-length in *vacuo*. But, by (2), this means that

$$\mu = \frac{\Lambda}{\lambda} > 1.$$

For electron waves, then, the metal is the denser medium optically. An electron ray incident obliquely is refracted towards the normal; when an electron ray moves out of a metal total reflection may occur (cf. below).

Whereas the velocity of the electron (group-velocity of the wave) is greater in the optically denser medium than outside, the phase-velocity, according to de Broglie's relation $ab = c^2$ [cf. eqn. (6), § 5, Chap. I], is *smaller* in the optically denser medium than outside. The distinction between group- and phase-velocity is essential for deciding the old controversy, corpuscular versus undulatory theory of light. According to Newton the velocity of the light corpuscles in the optically denser medium should be greater than outside (like our electron or group-velocity); but Foucault showed that the "velocity of light" in the optically denser medium is *less* (corresponding to our phase-velocity). We may express this by saying that the verdict on Newton's theory of optics was falsified by an ambiguous use of the word "velocity" and requires revision.

We now come to the numerical determination of the index of refraction. If eqn. (7) is integrated by assuming a plane wave, the wave-number [the quantity k in eqn. (13), § 1, Chap. I] becomes equal to the

root of the factor of ψ . Since the wave-length is inversely proportional to the wave-number, it at once follows from (7) that

$$\mu = \frac{\Lambda}{\lambda} = \sqrt{\frac{E + V_0}{E}} = \sqrt{1 + \frac{V_0}{E}} \quad (8)$$

The observations of Davisson and Germer make it possible to determine μ and hence V_0 empirically. It is assumed, of course, that the individual diffraction maxima are correctly interpreted, that is, allocated correctly to the reflecting net planes in the interior of the crystal. The method of allocation in the table differs from that which Davisson and Germer originally gave (in *Nature*) and agrees with that which they suggested as an alternative in their final publication (in the *Physical Review*).

TABLE 3.

E (Volt).	Λ (Å.).	λ (Å.).	μ .	V_0 (Volt).
54	1.67	1.49	1.12	18
106	1.19	1.18	1.06	11.5
160	0.97	0.92	1.05	14
188	0.89	0.85	1.04	15
190				
310	0.70	0.68	1.03	16
370	0.64	0.62	1.03	25

The first column contains the volt numbers for the incident electron ray; only a few representative results are taken out of the total number of experiments. The selectivity of the phenomenon and at the same time the space nature of the action of the lattice is expressed in the fact that every diffraction maximum occurred or attained its greatest intensity only at a definite velocity (that given in the first column). The second column gives the wave-length outside the crystal, calculated according to de Broglie's fundamental relation. The third column shows the wave-length in the interior of the crystal, which is different, calculated from Laue's theory in the sense of eqn. (3), or from Bragg's equation (which is identical with eqn. (3)). The values of μ in the fourth column are then obtained from the quotient of Λ and λ ; the average decrease of μ in the sequence from above downwards (increasing velocities) corresponds to eqn. (8), according to which μ approaches 1 more and more closely, the greater E becomes. Eqn. (8) at the same time shows how for a known value of μ the value of V_0 in the last column of the table is to be calculated. The mean value of V_0 from all the measurements of Davisson and Germer comes out approximately as

$$V_0 = 15 \text{ volts.}$$

The value of V_0 found in this way is instructive as giving us an insight into the nature of metals. For, according to the above remarks,

it gives us directly the work of escape which an electron must perform if it is to leave the metal, whereas the *Richardson effect* does not give this work itself, but only its excess over the work performed against the pressure which is exerted by the conduction electrons in the interior of the metals on the escaping electron.* The possibility of determining the work V_0 individually by electron experiments is clearly due to the electrons fired into the metal belonging to a range of velocity quite different from (higher than) the conduction electrons. This circumstance also shows that *Pauli's principle* (Chap. I, § 8) which governs the distribution of velocity of the conduction electrons does not restrict the motion of the rapid electrons fired into the metal.

We may enquire what place the law of refraction, which in eqn. (5) above, we simply took over from optics, is to have in the foundation of the theory of ψ -waves. Its origin in electrodynamics is, of course, no longer valid here: nor are we obliged to accept it. For the law of refraction is a kinematic relationship,—even in optics,—which is independent of the particular boundary conditions which obtain at the surface. It must always be fulfilled so long as the phase of the ψ -waves on both sides of the common surface is to be the same. For the propagation of this phase is determined by the value of $k \sin \theta$ (k = wave-number), and the equality of the propagation of this phase on both sides of the common surface already involves the law of refraction.

It is necessary to go into the question of the limiting conditions more closely in optics only if we wish to deduce the *Fresnel formula* for the coefficients of reflexion and refraction. We may also ask for the values of these coefficients in the case of electron waves, and we then also require definite boundary conditions. These are, as very general considerations show,

$$\psi_1 = \psi_2, \quad \left(\frac{\partial \psi}{\partial n} \right)_1 = \left(\frac{\partial \psi}{\partial n} \right)_2.$$

The fact that $\mu > 1$, for the transition vacuum \rightarrow metal and so $\mu' = 1/\mu < 1$, for the transition metal \rightarrow vacuum allows us to foresee *total reflexion* for such electrons as reach the surface of the metal at a sufficiently small angle from the interior. This is actually shown to occur in the experiments of Davisson and Germer: certain electron rays which are to be expected according to the uncorrected Laue theory are not observed outside; electron rays that emerge at very small angles (called "grazing beams" by Davisson and Germer) show that the index of refraction cannot be smaller than 1, as otherwise there would be a certain empty angular space in which no rays could be observed.

This sketch is, of course, only a first step to the complete theory of electron diffraction in the crystal. Just as the finer features of Röntgen ray interferences cannot be fully accounted for by the kinematic theory

* A. Sommerfeld, *Zeitschr. f. Phys.*, 47, 1 (1928); cf. in particular § 5.

of Laue but require the dynamic theory of Ewald,* so in the case of electron interferences in the crystal we must enlist a theory of wave-mechanics. We have first to expand the variable potential V in the metal lattice in a three-fold Fourier series and to assume the proper function ψ for the totality of electron waves in the corresponding form. The introduction of the index of refraction μ is justifiable in this way to a first approximation. In a second approximation we find a certain breadth for the maxima of reflexion and also a certain displacement of their position compared with that obtained by elementary theory. From the weakening which the electron waves experience through reflexion, we calculate the number of net planes which the electron wave can traverse; in the experiments of Davisson and Germer it amounts to between 10 and 100 layers; we have here disregarded the weakening of the electron wave by inelastic collisions with the metallic ions. Through the latter the number of layers traversed becomes further diminished; in the experiments of Davisson and Germer this number is *greatly* diminished because the velocities used by them are in a favourable region for the excitation process. Finally the mean value of the potential, that is, our work of escape V_0 , may also be determined purely theoretically, whereas we derived it empirically above from the observations. To do this it is necessary to use a wave-mechanical model—more or less similar to that of hydrogen—for the metallic atoms. The value of V_0 so found agrees very well with the above value of 15 volts for nickel, but yet depends on the number of electrons that are imagined detached from the individual metallic atom as conduction electrons. The possibility of determining this number from comparison with the observations gives a particularly interesting example of the application of experiments with electron rays, which is of importance for the theory of metals. Details of the more refined theory here indicated will be given in a comprehensive paper by H. Bethe.

B. Experiments by G. P. Thomson and others

As already mentioned in the introduction to this section, Thomson used thin foil or leaves of *micro-crystalline structure*: the material was of Au, Al, and celluloid; its thickness amounted to only about 300 atomic layers. Observations were made *behind* the leaves, the electron waves which passed through them being recorded *photographically*. The potential used was much greater than in the experiments of Davisson and Germer. In individual experiments it varied between 17.5 and 61.2 kV. By eqn. (1) the corresponding wave-length is 0.1 to 0.05 Å. This high value of the potential (magnitude of E) makes the index of refraction practically equal to 1 in Thomson's experiments. If we use a mean value

* P. P. Ewald, *Ann. d. Phys.*, 54, 519 (1918); cf. also *Handb. der Phys.*, 24, Springer, 1926.

of $E = 30,000$ volts and set $V_0 = 15$ volts, the value above obtained for nickel, we get from (8)

$$\mu = \sqrt{1 + \frac{15}{30000}} = 1.0002.$$

Thomson's photographs unmistakably have the character of Debye-Scherrer rings, especially the one obtained for gold leaf (Fig. 24). in which the grain was clearly sufficiently fine. In the case of aluminium (not reproduced here) the photograph is something between a Laue pattern and a Debye-Scherrer pattern; the micro-crystals seem in this



FIG. 24.

case to have been orientated with respect to the surface. The calculation of the diameter of the ring in the case of Au agrees within the limits of experimental error with those values which result from the known crystal dimensions of Au and from photographs with Röntgen rays. This opens up the possibility of analysing crystal structure with electron-waves instead of with Röntgen waves.

E. Rupp* has repeated Thomson's experiments with electrons of lower velocity. The most beautiful diffraction patterns with electrons of high velocity have been obtained by S. Kikuchi† by means of mica. According

to the thickness of the mica sheets either two-dimensional nets of the nature of patterns obtained from crossed lattices, that is, true Laue diagrams are produced, or else more complicated pictures traversed by light and dark lines, which correspond to the diffraction patterns of γ -rays previously described by Rutherford and Andrade.

§ 7. The Compton Effect

As was shown in I the change of wave-length of scattered Röntgen radiation discovered by A. H. Compton, and the electronic emission associated with it, is represented quite adequately by the corpuscular theory of light quanta. We should prefer still to regard this representation as the true causal treatment of the Compton effect. But it gives no answer to the following questions of intensity, which are of particular importance here. What is the relative intensity of the displaced line compared with the fixed line? How do these intensities depend on the angle of scattering? What is the state of polarisation of the displaced Compton line? What is the number of Compton electrons for the different angles of emission? The answers which Compton and Debye give to these questions vary (cf. *Atombau*, Note 2 in Chap. I, p. 56, and Note 3, p. 761). Here wave-mechanics again intervenes as the ideal statistical method.

* Ann. d. Phys., 85, 981 (1928).

† Japanese Journal of Physics, 5, 88, (1928).

A. The Wave-mechanical Equivalents of the Laws of Energy and Momentum

The fact that wave-mechanics can account for the Compton shift of wave-length seemed doubtful so long as the photo-electric effect had not been interpreted by wave-mechanics. But now that we have learned that the continuous spectrum is to be associated with the hyperbolic paths and electronic emission, the wave-mechanical interpretation of the Compton effect is indicated. We have only to look round for a phenomenon in which a primary wave simultaneously gives rise to two processes, of which one denotes the excitation of an energy level of the continuous spectrum, that is, electronic emission. We follow G. Wentzel* in recognising the phenomenon of "Smekal transitions" as being of this type, cf. § 3, D, Fig. 17a (Fig. 17b, in which the atom is supposed to be in an excited state initially does not come into question for the Compton effect). We set out these two phenomena according to the following scheme:—

TABLE 4.

Smekal transitions.	$h\nu$ energy element of the incident light.	$h\nu_{kl} = E_l - E_k$ excitation energy of the atom.	$\nu^* = \nu - \nu_{kl}$ non-coherent scattered radiation displaced towards the red.
Compton effect.	$E = h\nu$ energy element of the primary Röntgen radiation and of the undisplaced part of the Compton scattered radiation.	$h\nu_{kl} = E_l - E_k$ $E_l > 0$, energy in the continuous spectrum = energy of the recoil electron. $E_k < 0$, energy in the discontinuous spectrum, in particular in the ground state of the atom.	$\nu^* = \nu - \nu_{kl}$ Compton scattered radiation displaced towards the long-wave side.

Our scheme shows that the change of frequency ν_{kl} in the Compton effect arises from two amounts of energy: from the energy E_l of the electron emission and from the work of release ($-E_k$) of the emitted electron. At the same time our scheme shows that the wave-mechanical interpretation of the Compton effect takes over the one main support of the corpuscular theory, namely the law of conservation of energy—actually, the equating of the frequency is nothing other than the equating

* Zur Theorie des Compton-Effektes, I and II, Zeitschr. f. Phys., 43, 1 and 779 (1927). Cf. also the earlier works of W. Gordon, *ibid.*, 40, 117 (1926); E. Schrödinger, Ann. d. Phys., 82, 287 (1927); O. Klein, Zeitschr. f. Phys., 41, 407 (1927); G. Breit, Phys. Rev., 27, 362 (1926); P. A. M. Dirac, Proc. Roy. Soc., 111, 405, (1926). Wentzel's work denotes an advance as he considers *bound* electrons whereas the other authors quoted dealt with *free* electrons.

of the energy. The fact that the other main support of the corpuscular theory, the law of conservation of momentum, also comes into its own in the wave-mechanical interpretation will be seen presently.

We calculate the intensity of the scattered radiation in the Compton effect according to the formula (36) developed in § 3. In doing so we shall need to take into consideration only the first term of the right-hand side (vibration number $\nu^* = \nu - \nu_k$), since the second term ($\nu^* = \nu + \nu_k$), as above remarked, corresponds to an excited initial state. If we take eqn. 35 of § 3 into account, the formula is

$$M = \frac{hi}{4\pi c} \frac{e^2}{m} a e^{2\pi i \nu^* t} \sum_j \left(\frac{A_{ji}^* Q_{kj}}{E_i - E_j - h\nu} - \frac{A_{jk}^+ Q_{ij}^*}{E_k - E_j + h\nu} \right) \quad (1)$$

Following Wentzel we shall assume that the emitted electron is *weakly* bound compared with the energy $h\nu$ of the incident Röntgen radiation, that is, $|E_k| \ll h\nu$, and that the corresponding condition holds for all the energy-levels E_j that essentially occur in (1), $|E_j| \ll h\nu$, including the energy-levels in the continuous spectrum, in particular the energy E_i of the recoil electron. We may then replace the denominators in (1) by $\mp h\nu$ and obtain

$$M = \frac{1}{4\pi i} \frac{e^2}{m} \frac{a}{c\nu} e^{2\pi i \nu^* t} \sum_j (A_{ji}^* Q_{kj} + A_{jk}^+ Q_{ij}^*) \quad (2)$$

The summation over j may now be conveniently performed. Compare the analogous calculation of § 3. By (9) of § 3 the following relations hold [we have to choose the upper sign in (9) and to give A double indices]:

$$\left. \begin{aligned} \frac{\partial \psi_k}{\partial y} e^{-2\pi i \frac{x}{\lambda}} &= \sum A_{jk}^+ \psi_j \\ \frac{\partial \psi_i^*}{\partial y} e^{-2\pi i \frac{x}{\lambda}} &= \sum A_{ji}^- \psi_j^* \end{aligned} \right\} \quad (3)$$

In both equations the sum on the right-hand side denotes a summation over the discrete spectrum as well as an integration over the continuous spectrum. Further, we see from definition (34) of § 3 that the quantities Q_{kj} and Q_{ij}^* respectively, are expansion coefficients of a simple function namely:

$$\left. \begin{aligned} q e^{2\pi i \frac{x}{\lambda^*}} \psi_k &= \sum Q_{kj} \psi_j \\ q e^{2\pi i \frac{x}{\lambda^*}} \psi_i^* &= \sum Q_{ij}^* \psi_j^* \end{aligned} \right\} \quad (4)$$

By multiplying eqns. (3) and (4) together appropriately and integrating over the co-ordinate space we get

$$\int q e^{2\pi i (\frac{x}{\lambda^*} - \frac{x}{\lambda})} \left(\frac{\partial \psi_k}{\partial y} \psi_i^* + \frac{\partial \psi_i^*}{\partial y} \psi_k \right) d\tau = \sum (A_{jk}^+ Q_{ij}^* + A_{ji}^- Q_{kj})$$

For if we integrate the products of the sums that will originally have stood on the left of the last equation only the terms with $\psi_i \psi_j$ would remain, owing to the orthogonality of the ψ_j 's, and these terms would come to 1 owing to the normalising condition.

Our formula (2) for \mathbf{M} thereupon becomes

$$M = \frac{1}{4\pi i} \frac{e^2 a}{m_{cv}} e^{2\pi i \nu t} \int q \frac{\partial}{\partial y} (\psi_k \psi_l^*) e^{2\pi i (\frac{s}{\lambda^*} - \frac{x}{\lambda})} d\tau \quad (5)$$

The vector character of \mathbf{M} is contained in the factor $q = x, y, z$ under the integral sign. x is the direction of the primary Röntgen radiation, y the direction of the component of the primary electrical field as well as of the vector potential \mathbf{A} , corresponding to eqns. (1) and (1a), § 3, a is the amplitude of the vector potential, s denotes the optical path in the direction of the scattered radiation, reckoned from the centre of the

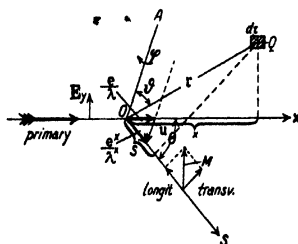


FIG. 25.

atom, as was explained in § 3, D; the direction of s is the direction in which the scattered radiation is observed.

In performing the integration in (5) we proceed in a manner similar to that in § 5, Fig. 21. Let \mathbf{e} be the unit vector in the direction of the x -axis (primary radiation), \mathbf{e}^* the unit vector in the direction of the optical path s (scattered radiation), Q the point of integration, situated in the element of integration $d\mathbf{r}$, and $\mathbf{r} = OQ$. We then have

$$x = (\theta, \mathbf{r}), \quad s = (\theta^*, \mathbf{r}).$$

and hence

$$\frac{s}{\lambda^*} - \frac{x}{\lambda} = \left(\frac{\theta^*}{\lambda^*} - \frac{\theta}{\lambda}, \tau \right) \quad (6)$$

Following Wentzel, we introduce a co-ordinate system r, θ, ϕ , which has the vector u for the polar axis (Fig. 25), where

$$u = \frac{\theta^*}{\lambda^*} - \frac{\theta}{\lambda} \quad (7)$$

216 Chapter II. Perturbations. Diffraction Problems

θ denotes the angle between $\mathbf{r} = \text{OQ}$ and this axis OA , ϕ the angle described about this axis. We may then write eqn. (6), according to the meaning of the scalar product:

$$\frac{s}{\lambda^*} - \frac{x}{\lambda} = |\mathbf{u}| r \cos \theta \quad . \quad . \quad . \quad (8)$$

We may reckon the proper function ψ_k as being spherically symmetrical, since it is the ground state of the atom; in accordance with eqn. (2) in § 7, Chap. I, ψ_i , being a proper function in the continuous spectrum will depend on the angles in a manner denoted by $f(\theta, \phi)$. If we use the hydrogen formulæ of Table I, § 7, Chap. I, on p. 71, and the asymptotic representation in eqn. (24), p. 67, we get

$$\left. \begin{aligned} \psi_k &= \dots e^{-k_1 r}, & \psi_i^* &= \dots f(\theta, \phi) \frac{1}{r} e^{\pm i k_2 r} \\ k_1 &= \frac{1}{a}, & k_2 &= \frac{2\pi}{h} \sqrt{2mE_i} \end{aligned} \right\} \quad . \quad . \quad (9)$$

The result

$$\frac{\partial}{\partial y} \psi_k \psi_i^* \sim \frac{1}{r} \frac{\partial r}{\partial y} e^{(-k_1 \pm i k_2)r} f(\theta, \phi) \quad . \quad . \quad . \quad (9a)$$

follows if we perform the differentiation asymptotically and strike out the unimportant factor $(-k \pm i k_2)$; for the terms with $\frac{\partial \theta}{\partial y}, \frac{\partial \phi}{\partial y}$ vanish more rapidly with $\frac{1}{r}$ than the term written down. Since the co-ordinate q is proportional to r and since

$$\frac{\partial r}{\partial y} = \frac{y}{r}$$

is independent of r , we see that

$$\frac{q}{r} \frac{\partial r}{\partial y} = \frac{q}{r} \frac{y}{r} = f_1(\theta, \phi) \quad . \quad . \quad . \quad (10)$$

is also independent of r . The integral to be integrated in (5) with respect to r now runs, in consequence of (8), (9a) and 10,

$$\int_0^\infty r^2 dr e^{(-k_1 \pm i k_2 + 2\pi i |\mathbf{u}| \cos \theta)r},$$

and comes out as

$$\frac{2}{(-k_1 \pm i k_2 + 2\pi i |\mathbf{u}| \cos \theta)^3} \quad . \quad . \quad . \quad (11)$$

Next, the integration with respect to θ and ϕ is to be performed, namely, by (5), (10) and (11):

$$\iint \frac{F(\theta, \phi) \sin \theta \, d\theta d\phi}{(a + b \cos \theta)^3} \dots \left\{ F(\theta, \phi) = f(\theta, \phi) \cdot f(\theta, \phi) \right\} \quad (12)$$

The function $F(\theta, \phi)$ may be expanded in surface harmonics $P_l^m(\cos \theta) e^{im\phi}$; if we straightway integrate with respect to ϕ and set $x = \cos \theta$, (12) resolves into a sum over l of integrals of the form

$$\int_{-1}^{+1} \frac{P_l^0(x) dx}{(a + bx)^2} \quad (13)$$

We integrate by parts and retain only the term with the highest negative power of $a + bx$, which is alone decisive for us, as we shall presently see. This term is

$$\text{Const.} \frac{1}{(a \pm b)^2}$$

and so, by (12), is proportional to

$$(-k_1 \pm ik_2 \pm 2\pi i |u|)^{-2} \quad (14)$$

Here we can neglect k_1 in comparison with k_2 , if we wish to take into consideration only recoil electrons of a fairly large observable velocity. For, by (9), we have

$$\left(\frac{k_2}{k_1}\right)^2 = \frac{8\pi^2 m a^2}{h^2} E_i$$

thus with $a = h^2/4\pi^2 m e^2$ (radius of the hydrogen atom) and

$$|E_k| = 2\pi^2 m e^4 / h^2$$

(work of ionisation for hydrogen = amount of energy in the ground state),

$$\left(\frac{k_2}{k_1}\right)^2 = \frac{E_i}{|E_k|} \quad (15)$$

Thus so long as the energy E_i of the recoil electron is many times greater than the work of release $|E_k|$, k_2 will also be many times greater than k_1 and the right-hand side of (14) reduces to

$$-\frac{h^2}{4\pi^2} \left(\sqrt{2mE_i} \pm h \left[\frac{e^*}{\lambda^*} - \frac{e}{\lambda} \right] \right)^{-2} \quad (16)$$

Now $\sqrt{2mE_i}$, with $E_i = mv^2/2$, is the momentum mv of the recoil electron. Thus the other terms within the same bracket are also momentum quantities, namely

$$\left. \begin{aligned} \frac{h}{\lambda} &= \frac{h\nu}{c} = p = \text{momentum of the primary radiation} \\ e^* \frac{h}{\lambda^*} &= e^* \frac{h\nu^*}{c} = p^* = \text{momentum of the secondary radiation} \end{aligned} \right\} \quad (16a)$$

If we use the lower sign in (16) and set the expression in brackets equal to zero, we get the law of conservation of momentum: the momentum of the recoil electron is made up of the vector difference between the primary and the secondary radiation:

$$mv = |\vec{p} - \vec{p}^*| \quad (17)$$

Equating to zero in this way corresponds to a sharp maximum in the expression (14) and hence also in the expression (5) for the moment and in the intensity of the scattered radiation. Thus the secondary momentum p^* in eqn. (17) is to be understood in particular as that momentum which corresponds to the *maximum* of the secondary spectrum emitted. Hence the Compton-Debye formula (*Atombau*, Chap. I, p. 56, $d\lambda = 2\lambda_0 \sin^2 \theta/2$) holds rigorously for the difference of wave-length which this maximum has as compared with the primary wave-length.

In (17) we have deduced only a part of the law of momentum by wave-mechanics, since we calculated only the *magnitude* of the momentum mv and not its *direction* from the radiation-momenta \vec{p} and \vec{p}^* . This signifies a certain gap in our representation. It is true that this part of the law of momentum, combined with the energy law, suffices for determining the Compton change of wave-length $\Delta\lambda$ (cf. *Atombau*, Note 3, eqns. (1) and (3), or (1') and (3')), where we used only the formulæ just now re-discovered by wave-mechanics, namely the energy law and that part of the law of momentum, which refers to the *magnitude* of the momentum, whereas we did not in our earlier calculation require the fact that the directions of the two momenta are the same.

But observations* make it appear probable that the momentum law is valid in its full vectorial sense in the Compton effect.

This is also shown by those theoretical treatments (Gordon, Schrödinger, etc., cf. footnote, p. 212), which deal with *free* electrons. In them the momentum law is obeyed in a sharp sense. It is easy to understand that in our treatment and Wentzel's, which deal with bound electrons, the momentum law comes out unsharp. For the electron can derive a part of its momentum from the atom and this part is, like all atomic happenings, only statistically determined. To complete our treatment we should have to show that this statistical unsharpness of the momentum-law decreases to zero as the binding energy decreases.

B. Width of the Displaced Compton Line. Polarisation. Intensity

The transition from (14) to (16) shows that the maximum of the displaced intensity becomes increasingly sharper as k_1 , that is, the work of release J becomes smaller. If we cannot neglect k_1 , the maximum of the expression (14) does not become infinitely sharp, but spreads out similarly to the well-known conditions in the case of a resonance-maximum with a damping term. *A pronounced Compton line is to be expected only in the case of outer electrons, whose energy is of the order of the work of release**

* A. H. Compton and A. W. Simon, *Phys. Rev.*, **25**, 306 (1925). Observations by F. Kirschner, cf. below, can be interpreted in the same sense. Important observations by Bothe and Geiger exhibit the *simultaneous* occurrence of recoil electrons and scattered radiation, but give no information about the dependence of these two processes on direction. It is, however, reported that the most recent results obtained in Compton's laboratory do not confirm the allocation of directions given by Compton and Simon.

(ionisation energy) for the hydrogen atom ($|E_k| = 13.5$ volts); the more tightly bound electrons of the K-, L-, . . . shells of heavy atoms give only a *blurred* impress, no pronounced line. This is in agreement with the view taken by Compton* that the recoil electrons are (nearly) free electrons, but the photo-electrons are bound electrons. Of course, it is not the absolute value of the ionisation potential which is significant but its magnitude relative to that of the primary light-quantum. In the case of hard primary radiation the displaced line becomes sharper, according to theory, than in the case of soft primary radiation,† the scattering bodies being assumed to be the same [cf. the calculation of $E_l/|E_k|$ in eqn. (28) at the end of this section].

It would be false to assume that the sharpness of the Compton line would have to decrease with increasing atomic weight of the scattering body. For the strength with which the *outer* electrons are bound is no wise proportional to the atomic weight of the elements, but remains constant on the average in the periodic table. (In the case of the alkalis, indeed, it decreases in the row Li to Cs, cf. I, Table 36, p. 345.) Actually, observations by Woo‡ made with an ionisation chamber and dealing with a large number of different scattering bodies (from Li to Cu), and photographs taken by P. A. Ross,§ show that there is no decrease in the sharpness of the lines with increasing atomic weight.

The *polarisation* of the Compton radiation can be best found by transforming the integral (5) by integration by parts with respect to y . For $q = y$ there thus results a different form of dependence than for $q = x$ or $q = z$, namely,

$$M_x = N_x, \quad M_y = M + N_y, \quad M_z = N_z \quad . \quad . \quad (18)$$

M here arises from the differentiation of the factor q under the integral sign in (5) with respect to y . If we omit the factor in front of the integral sign, we have

$$M = - \int \psi_k \psi_l^* e^{2\pi i \left(\frac{t}{\lambda^*} - \frac{x}{\lambda} \right)} d\tau \quad . \quad . \quad . \quad (19)$$

On the other hand, N results from differentiating the exponent of e under the integral sign in (5) with respect to y ; we get the following factor

$$\frac{2\pi i}{\lambda^*} \frac{\partial s}{\partial y} = \frac{2\pi i}{\lambda^*} \cos(s, y).$$

For N we get in this way ($q = x, y, z$)

$$N_y = - \frac{2\pi i}{\lambda^*} \cos(s, y) \int q \psi_k \psi_l^* e^{2\pi i \left(\frac{t}{\lambda^*} - \frac{x}{\lambda} \right)} d\tau \quad . \quad . \quad (20)$$

* Phys. Rev., **31**, 59 (1928).

† Wentzel (*loc. cit.*, p. 193) gives experimental indications of this.

‡ A. H. Compton and Y. H. Woo, Proc. Nat. Acad., **10**, 271 (1924); cf. also A. H. Compton, *X-rays and Electrons*, New York, 1926, p. 268.

§ Proc. Nat. Acad., **10**, 804 (1924), and A. H. Compton (*loc. cit.*, p. 269).

If we make a very rough estimate we may say that N is to M as \bar{q} is to λ^* , where \bar{q} stands for a mean of the radius vector q taken over the volume of the atom (a mean radius of the retarded charge-density ρ_{el}). From this we infer that in the case of Röntgen radiation that is not too hard $|N| \ll |M|$ and hence also $|M_s, M_z| \ll |M_y|$ is valid. That is to say: *the scattered radiation in the Compton effect will show essentially the same polarisation as the primary Röntgen radiation*, so far as this is compatible with the transversality of the scattered radiation. To determine the polarisation of the scattered radiation, we must divide the exciting momentum into two components (cf. Fig. 25, p. 214), a longitudinal component in the direction of emission s and a transverse component perpendicular to it. By the general laws of electrodynamics only this transverse component of the moment gives radiation in the s -direction. If we observe, in particular, in the direction $\Theta = \pi/2$, namely along the y -axis, and if, as we have hitherto assumed, the primary radiation is polarised in this direction, the transverse component of M and hence also the intensity of the displaced Compton line is zero. The same, of course, holds for the undisplaced line of the scattered light. Expressed in other words: *the conditions of polarisation in the displaced and undisplaced lines are the same*. If the primary radiation is not polarised but unpolarised, so that besides E_y there is a primary component E_x , the Compton line, just like the ordinary scattered line, will be completely polarised in the z -direction when observed in the y -direction, and *vice versa*.

Experimental proof of these facts is given by very beautiful researches by H. Kallmann and H. Mark,* who compare the intensity of the displaced line with that of the undisplaced line in the neighbourhood of the angle of scattering $\Theta = \pi/2$, after both had been reflected by a crystal. The ratio of the two was found to be constant, that is, their polarisation was the same. The method is modelled on the classical arrangement of Barkla (cf. I, Chap. I, p. 26). An indirect proof of the polarisation of the displaced Compton line is given by observations of the corresponding recoil-electrons in a Wilson-chamber, carried out by F. Kirohner.†

The observations of Kallmann and Mark were made with fairly soft Röntgen rays (K -series of zinc). In this case we are justified in eqn. (18) in neglecting N compared with M . In the case of harder rays N must of course be taken into consideration; the angle of complete polarisation (vanishing intensity of the parallel component) also then, however, appears to lie ‡ at $\pi/2$ in contradiction of earlier semi-classical arguments involving the so-called "substitute electron" (*Ersatz-Elektron*).

* Zeitschr. f. Phys., **36**, 120 (1926); cf. also the comprehensive survey by the same authors in Vol. V, of the "Ergebnisse der exakten Naturwissenschaften," Springer, 1928.

† Phys. Zeitschr., **27**, 885 (1926).

‡ This was first shown theoretically in the paper by Dirac quoted on p. 212 and then confirmed experimentally under Compton's supervision by C. S. Barnett and J. A. Bearden, Phys. Rev., **29**, 352 (1927).

Passing on to questions of intensity we next follow Wentzel and calculate the total scattered radiation, that is, the sum of the intensities J , which are emitted in transitions from the ground level k to all possible levels l of the continuous and of the discontinuous spectrum. In doing so, we use the simplified value (19) of the moment, which we shall now call M_{kl} , and, as in (19), we shall leave out the factor in front of the integral sign. We then have

$$\sum J = \sum_l |M_{kl}|^2 = \sum_l \int \psi_k \psi_l^* e^{i\eta} d\tau \cdot \int \psi_k^* \psi_l e^{-i\eta} d\tau \quad (21)$$

with the abbreviation

$$\eta = 2\pi \left(\frac{s}{\lambda^*} - \frac{x}{\lambda} \right).$$

The value of (21) comes out simply as 1. To prove this we proceed in a manner similar to that used in eqns. (3). We expand the two functions

$$f = \psi_k e^{+i\eta} \quad \text{and} \quad g = \psi_k^* e^{-i\eta} \quad (22)$$

according to the complete system of proper functions ψ_l :

$$f = \sum A_{kl} \psi_l, \quad g = \sum B_{kl} \psi_l^* \quad (23)$$

and find from this by means of multiplication and integration over the region of co-ordinates, paying due regard to the orthogonality and normalisation of the proper functions,

$$\int f g d\tau = \sum A_{kl} B_{kl} \quad (24)$$

In our case the left-hand side becomes, on account of (22),

$$\int \psi_k \psi_k^* d\tau = 1;$$

on the other hand, the right-hand side of (24) becomes identical with the right-hand side of (21). For by (23) we have

$$A_{kl} = \int f \psi_l^* d\tau = \int \psi_k \psi_l^* e^{i\eta} d\tau,$$

$$B_{kl} = \int g \psi_l d\tau = \int \psi_k^* \psi_l e^{-i\eta} d\tau.$$

Eqn. (21) together with (24) thus actually expresses that the total intensity becomes $\sum J = 1$, which is what was asserted.

We also remark that eqn. (24) is a generalised form of the so-called "completeness relation" ("*Vollständigkeits-Relation*"), which we touched on in § 1 A, cf. the footnote on p. 147.

In the case of very long waves for which $\eta = 0$ we clearly have $M_{kl} = 0$ or 1, according as $l \neq k$ or $= k$. The sum (21) then consists of

only the one term 1. In this case it is obvious that the intensity-sum here calculated merges into the value of the intensity of the scattered radiation, as obtained by classical calculations. But since the value 1 of the sum, as we have shown, is generally valid [so long as it is permissible for us to neglect the component N , eqn. (18)], the result that *the emitted radiation, summed up over all the energy levels l retains the classical value independently of the wave-length of the primary radiation* remains unaffected. This can be tested by taking into account in (5) the factor in front of the integral sign, which was omitted in (19) as well as by considering the meaning of a [amplitude of the vector potential, eqn. (1) of § 3]. This method, besides showing that the intensity is independent of the wave-length, also seems to show that it is independent of the angle of scattering, in opposition to the classical theory of scattering. We must not lose sight of the fact, however, that only the component of M *transverse* to the direction of scattering Θ contributes to the emission, so that actually in forming the components the factor $\sin \Theta$ and, in the intensity, the factor $\sin^2 \Theta$ occurs, as should be the case according to the classical formula for scattering.

For very short waves our calculation is insufficient; as already emphasised in dealing with polarisation, the component N of the moment may then no longer be neglected in comparison with M . That is why the classical formula of scattering fails for γ -rays; the scattering coefficient approaches the value zero (cf. *Atombau*, Note 3, p. 761). The reason for this is that in the integration over the atom the zones of different sign for $\cos \gamma$ cancel each other owing to interference. This was calculated in greater detail for the undisplaced line by J. A. Stratton.*

In reckoning up the sum of all the transitions $k \rightarrow l$ in (21) we also included the "transition" $k \rightarrow k$. This corresponds to ordinary scattering, that is, to the undisplaced line in the Compton effect or, in the sense of § 3, to ordinary dispersion not modified by Smekal transitions. So we determine the intensity of the undisplaced line to the same approximation as before from the "diagonal matrix element":

$$M_{kk} = \int \psi_k \psi_k^* e^{i\gamma r} d\tau \quad . \quad . \quad . \quad (25)$$

If we insert the value of ψ_k from (9) and add the normalising factor (cf. Table I, p. 71), we have

$$\psi_k = \sqrt{\frac{k^3}{\pi}} e^{-k_1 r} \quad .$$

Hence we must form

$$M_{kk} = \frac{k^3}{\pi} \int e^{-2k_1 r + i\gamma r} d\tau.$$

* *Helvetica physica acta*, 1, 47 (1928).

We use the earlier abbreviation (8) and the same co-ordinate system r, θ, ϕ as in Fig. 25. We get

$$\begin{aligned} M_{kk} &= 2k_1^3 \int_0^\pi \sin \theta \, d\theta \int_0^\infty e^{-2k_1 + 2\pi i |u| \cos \theta} r^2 \, dr \\ &= 4k_1^3 \int_0^\pi \frac{\sin \theta \, d\theta}{(-2k_1 + 2\pi i |u| \cos \theta)^3} \\ &= 4k_1^3 \int_{-1}^{+1} \frac{dx}{(a + bx)^3} \dots \begin{cases} x = \cos \theta \\ a = -2k_1 \\ b = 2\pi i |u| \end{cases} \\ &= \frac{2k_1^3}{b} \left(\frac{1}{(a-b)^2} - \frac{1}{(a+b)^2} \right) = - \left(1 + \frac{\pi^2 |u|^2}{k_1^2} \right)^{-2}. \end{aligned}$$

From eqn. (14) we can see that for the maximum of the displaced line $2\pi |u| = k_1$. So we can write

$$|M_{kk}| = \left(1 + \frac{k_2^2}{4k_1^2} \right)^{-2},$$

or, on account of (15)

$$|M_{kk}| = \left(1 + \frac{E_l}{4|E_k|} \right)^{-2}. \quad (26)$$

The intensity of the undisplaced line decreases with the ratio of the binding-energy $|E_k|$ to the recoil-energy E_l (more exactly, to the recoil-energy that corresponds to the maximum of the displaced line).

This result, which according to the earlier meaning of $|E_k|$ was first derived for the hydrogen-electron, can also be applied to the various electrons of the other atoms.

From (26) we at once get the ratio of the intensities of the displaced and the undisplaced lines. For since the intensity is generally measured by the square of the moment and since the sum of the intensities is equal to 1, we get for the ratio in question

$$\frac{1 - M_{kk}^2}{M_{kk}^2} = \left(1 + \frac{E_l}{4|E_k|} \right)^4 - 1. \quad (27)$$

To carry out the calculation further we get from the Table on p. 213,

$$E_l - E_k = h(\nu - \nu^*), \quad \frac{E_l}{|E_k|} = -1 + \frac{\nu - \nu^*}{\nu_k},$$

if $\nu_k = |E_k|/h$ denotes the ionising frequency in the state E_k . Further we set $\nu_k = c/\lambda_k$ and express the Compton change of frequency $\nu - \nu^* = -\Delta\nu$ by the corresponding change of wave-length $\Delta\lambda$:

$$\nu - \nu^* = c \frac{\Delta\lambda}{\lambda^2} = c \frac{2\lambda_0}{\lambda^2} \sin^2 \frac{\Theta}{2}.$$

Accordingly

$$\frac{E_l}{|E_k|} = -1 + \frac{2\lambda_0\lambda_k}{\lambda^2} \sin^2 \frac{\Theta}{2}. \quad (28)$$

and
$$\frac{1 - M_{kk}^2}{M_{kk}^2} = \left(\frac{3}{4} + \frac{\lambda_0 \lambda_k}{2\lambda^2} \sin^2 \frac{\Theta}{2} \right)^4 - 1 \quad (29)$$

In this we substitute the values

$$\begin{aligned} \lambda_0 &= 24 \cdot 10^{-3} \text{ \AA. (from Atombau, p. 57),} \\ \lambda_k &= 10^3 \text{ \AA. (approximate limit of the Lyman series),} \\ \lambda &= 1 \cdot 5 \text{ \AA. (K-radiation of copper).} \end{aligned}$$

We then get for $\Theta = \pi/2$

$$\frac{\lambda_0 \lambda_k}{2\lambda^2} \sin^2 \frac{\Theta}{2} \sim 2 \cdot 7, \quad \frac{1 - M_{kk}^2}{M_{kk}^2} \sim 140.$$

Thus most of the intensity goes into the displaced line. This is the more so the harder the primary radiation and the greater the angle of scattering.

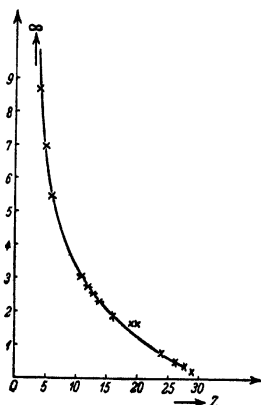


FIG. 26.—Intensity-ratio R of the displaced to the undisplaced line for $\lambda = 0 \cdot 709 \text{ \AA}$ and angle of scattering $\Theta = 120^\circ$ as a function of the atomic number Z of the scattering body (drawn according to the numerical data of Woo).

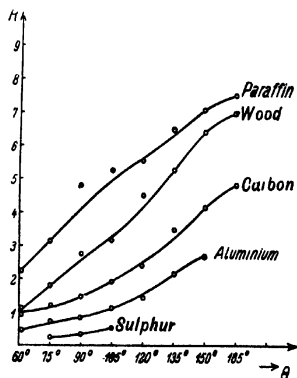


FIG. 27.—Intensity-ratio R of the displaced to the undisplaced line for $\lambda = 0 \cdot 709 \text{ \AA}$ (K_α of molybdenum) as a function of the angle of scattering Θ (according to Woo).

But if we consider the electrons of the inner atomic shells, the work of release in their case is considerably greater and λ_k correspondingly smaller. If we calculate with a λ_k ten times smaller then we get for our intensity ratio, other assumptions remaining the same, the value $0 \cdot 08$, that is, *the principal part of the intensity goes into the undisplaced line*. Moreover, at the same time the displaced line becomes blurred, as we saw at the beginning of this section; it thus eludes our observation, both on account of its feeble intensity and on account of its diffuse character. This result, too, agrees with Compton's view that the displaced line and

the recoil-phenomenon essentially arise from the free (loosely bound) electrons. The scattering from the more tightly bound electrons contributes appreciably only to the undisplaced line.

The fact that for a very small λ_k , for example, for the K-electrons of heavy elements we get a negative value from eqn. (29) for our intensity-ratio is obviously due to our allowing certain terms to be neglected.

A very beautiful confirmation of these results is shown in Figs. 26 and 27, which have been drawn * in accordance with the results of Woo (of the paper quoted on p. 218). Fig. 26 gives the intensity ratio R of the disturbed and undisplaced lines as a function of the atomic number Z of the scattering body from Li to Cu. Fig. 27 shows the same ratio as a function of the scattering angle Θ for a few chosen substances. The primary wave-length is everywhere the same: $\lambda = 0.709 \text{ \AA.}$ (K_α of molybdenum). In the case of Li ($Z = 3$) the whole intensity goes into the displaced line, so that in Fig. 26 the ratio R becomes ∞ ; as the number of tightly bound electrons increases, the undisplaced line increases in intensity and the ratio R tends to zero. The increase of R with increasing Θ in Fig. 27 also agrees with our formula (29).

§ 8. The Helium Problem

A short summary of the older suggestions about the structure of the helium model was given in *Atombau*, Chap. III, p. 198 (see also I, p. 96). These suggestions all failed as they gave results in direct contradiction to experiment; among these contradictions we must count the existence of a magnetic moment in the ground-state of helium, which presented itself in several of these models. The author's suggestion (*Atombau*, Fig. 41, p. 203) is, indeed, free from this objection, but has likewise lost its purpose, as the question as to how a collision between plane equivalent 1_1 -orbits is to be avoided no longer occurs, since the conception of orbits has been eliminated by wave-mechanics.

The beauty of the *wave-mechanical model of helium* is that it renders unnecessary any visual pictures whatsoever, but is contained in the general wave-equation for two electrons:

$$\Delta\psi + \frac{8\pi^2m}{h^2}(E - V)\psi = 0 \left\{ \begin{array}{l} \Delta = \Delta_1 + \Delta_2 \\ V = -\frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}} \end{array} \right\} \quad (1)$$

Here Δ denotes the differential parameter in the six-dimensional configuration space of the two electrons (1) and (2), and it is composed additively of the two three-dimensional differential expressions Δ_1 and Δ_2 . The potential energy has been written down for a Z -fold charged nucleus; $Z = 2$ denotes He; $Z = 3$ denotes Li^+ ; $Z = 4$, Be^{++} . r_{12} is the dis-

* Our thanks are due to Mr. F. Kirchner for allowing us to use them here.

tance of the two electrons from each other; r_1, r_2 are their respective distances from the nucleus.

We may regard it as agreed that the new "model" is in accordance with the facts. Heisenberg* was able to show that it gives the correct structure of the helium spectrum qualitatively, namely, that it accounts for the division of the spectrum into the two non-combining partial systems *ortho*- and *par*-helium: he was also able to calculate the mutual position of the two partial systems approximately correctly. Further, by using the idea of the spinning electron, he determined the *fine-structure* of the *ortho*-system as abnormal triplet, which was promptly confirmed by the very careful experimental investigations of W. V. Houston† and G. Hansen.‡ Finally numerical calculations (cf. Chap. I, § 9, p. 112) were possible which allowed the energy in the ground state of helium and hence the ionisation potential of helium to be deduced; this also now agrees exactly with experiment. An analytical calculation by Unsöld§ along the lines suggested by Heisenberg approaches the correct result as closely as may be expected from the bad convergence of the perturbation method for the ground-state.

A. First Approximation and Scheme of the Perturbation Theory. General Structure of the Helium Spectrum

To a first approximation we neglect the inter-action term e^2/r_{12} and can then separate eqn. (1). We set

$$E = E_1 + E_2, \quad \psi = \psi(1) \cdot \psi(2)$$

in which we arbitrarily denote the one electron by 1 and the other by 2. The separated equations are

$$\Delta_1 \psi(1) + \frac{8\pi^2 m}{h^2} \left(E_1 - \lambda + \frac{Ze^2}{r_1} \right) \psi(1) = 0. \quad (2)$$

$$\Delta_2 \psi(2) + \frac{8\pi^2 m}{h^2} \left(E_2 + \lambda + \frac{Ze^2}{r_2} \right) \psi(2) = 0. \quad (3)$$

λ is the separation parameter which is introduced according to the method of eqn. (1a, b), § 2, Chap. I; that is, it is a quantity *independent* of all the co-ordinates. We shall have to make it equal to zero in section C.

Eqns. (2) and (3) differ from the simple wave-equation for the hydrogen atom only in formal respects, in that the proper-value parameter is called $E_1 - \lambda$ and $E_2 + \lambda$ respectively. Using the known proper values and proper functions of the Kepler problem we thus have, if we integrate

* Über die Spektren von Atomsystemen mit zwei Elektronen, *Zeitschr. f. Phys.*, **39**, 499 (1926). Cf. also the introductory paper, *ibid.*, **33**, 411 (1926), and the generalisation to problems of many bodies, *ibid.*, **41**, 289 (1927).

† *Phys. Rev.*, **29**, 749 (1927).

‡ Kindly communicated personally.

§ *Ann. d. Phys.* (cf. the reference on p. 85 of § 8, Chap. I).

(2) by means of the proper function ψ_1 for $n = 1$ and (3) by the proper function ψ_n for $n > 1$,

$$E_1 - \lambda = -\frac{R\hbar Z^2}{1^2}, \quad E_2 + \lambda = -\frac{R\hbar Z^2}{n^2}. \quad (4)$$

Our solution of the helium problem to a first approximation thus runs

$$u = \psi_1(1)\psi_n(2) \quad (5)$$

It satisfies the differential equation of the first approximation,

$$L(u) = \Delta u + \frac{8\pi^2 m}{\hbar^2} \left(E_1 + E_2 + \frac{Ze^2}{r_1} + \frac{Ze^2}{r_2} \right) u = 0. \quad (6)$$

But the same differential equation is also solved by the assumption

$$v = \psi_1(2)\psi_n(1) \quad (7)$$

In the configuration space of our six co-ordinates (7) and (5) are two *quite distinct* functional relationships except when $n = 1$, which we have excluded. But their proper values are *exactly equal*, namely, to $E_2 + E_1$. This is therefore a case of “*degeneracy*”; we shall call it “*exchange-degeneracy*” (*Austausch-Entartung*). Heisenberg also uses the term “*resonance-degeneracy*” (*Resonanz-Entartung*), because the absolute equality of mass and binding of the two electrons, which leads in ordinary mechanics to resonance phenomena, is also a condition for the present case of degeneracy.*

In the case of exchange-degeneracy the conditions are exactly as in that of the square membrane, Fig. 14 in § 1, in which all proper values $k_{m,n}$ with $m \neq n$ were *two-fold*, and only those with $m = n$ were *single*. In the case of helium the only single proper value that comes into consideration practically is that of the ground state, since, as we know, the inner electron is in the K-shell, so that the condition for non-degeneracy reduces to $n = 1$.

We have to bear in mind the fact of degeneracy when we now pass on to correct our first approximation by means of a perturbation calculation. Since neither of the solutions u and v has an advantage over the other we form a combination of them,

$$w = \alpha u + \beta v \quad (8)$$

with arbitrary constant coefficients α and β and set

$$\psi = w + \phi, \quad E = E_1 + E_2 + \eta \quad (9)$$

* Besides this exchange degeneracy we also have the degeneracy in angular co-ordinates known to us from hydrogen, since to the given energy value (given principal quantum number n) there belong n^2 different proper functions which differ in regard to the quantum numbers l and m . Of these two cases of degeneracy, namely in l and m , we regard the m -degeneracy as eliminated by means of a magnetic field. The degeneracy in l is eliminated owing to the disturbing interaction between the two electrons; it is n -fold, as may easily be seen, cf. eqn. (10a), § 7, Chap. I. We shall therefore disregard this degeneracy in the sequel as it is of no account practically. To prove this we should have to resolve this degeneracy by means of a perturbation-calculation (with a determinant of the n th order). This would lead to energy levels extremely near to each other.

where ϕ and η are small quantities, whose products with the perturbation-member ϵ^2/r_{12} may be neglected. For the differential equation in ϕ we get from (1) and (6)

$$\Delta\phi + \frac{8\pi^2m}{h^2}\left(E_1 + E_2 + \frac{Ze^2}{r_1} + \frac{Ze^2}{r_2}\right)\phi = (s - \epsilon)w \quad (10)$$

On the right hand we have used the abbreviations

$$s = \frac{8\pi^2m}{h^2} \frac{e^2}{r_{12}}, \quad \epsilon = \frac{8\pi^2m}{h^2} \eta \quad (11)$$

in order to fit in with the working of § 1, B, which now comes into action for the first time.

As in § 1, B, we say that in order that the "non-homogeneous" eqn. (10) should have a solution at all, its right-hand side must be "orthogonal" to the solutions u and v of the associated "homogeneous" eqn. (6). Hence the conditions

$$\int (s - \epsilon)wud\tau = 0 \quad \text{and} \quad \int (s - \epsilon)wvd\tau = 0 \quad (11a)$$

Here $d\tau = d\tau_1 \cdot d\tau_2$ is the 6-dimensional volume element of all the degrees of freedom and it subdivides into the two 3-dimensional volume-elements $d\tau_1$ and $d\tau_2$. Inserting the value (8) of w , we get two conditions for α and β :

$$\left. \begin{aligned} \alpha \int (s - \epsilon)u^2 d\tau + \beta \int (s - \epsilon)uv d\tau &= 0 \\ \alpha \int (s - \epsilon)uv d\tau + \beta \int (s - \epsilon)v^2 d\tau &= 0 \end{aligned} \right\} \quad (12)$$

Let the u 's and v 's be normalised to 1, so that

$$\int u^2 d\tau = \int v^2 d\tau = 1.$$

Further, they are orthogonal to each other. For

$$0 = \int uv d\tau = \int \psi_1(1)\psi_n(1)d\tau_1 \cdot \int \psi_1(2)\psi_n(2)d\tau_2$$

as both the integral in $d\tau_1$ and that in $d\tau_2$ vanish.

If we link up with p. 152 by introducing the abbreviations

$$\epsilon_{11} = \int su^2 d\tau, \quad \epsilon_{12} = \epsilon_{21} = \int suv d\tau, \quad \epsilon_{22} = \int sv^2 d\tau \quad (13)$$

we may write, bearing in mind that ϵ , being constant, can be placed outside the integral, in place of (12),

$$\left. \begin{aligned} \alpha(\epsilon_{11} - \epsilon) + \beta\epsilon_{12} &= 0 \\ \alpha\epsilon_{12} + \beta(\epsilon_{22} - \epsilon) &= 0 \end{aligned} \right\} \quad (12a)$$

Eliminating α and β we get a quadratic equation for ϵ :

$$\begin{vmatrix} \epsilon_{11} - \epsilon & \epsilon_{12} \\ \epsilon_{12} & \epsilon_{22} - \epsilon \end{vmatrix} = 0 \quad (14)$$

Now v arises from u and hence also ϵ_{22} from ϵ_{11} by exchanging the two electrons. But since the integration with respect to dr considers both electrons equally, and since also s [eqn. (11)] is symmetrical in both electrons, we get

$$\epsilon_{22} = \epsilon_{11}.$$

Our quadratic eqn. (14) thus becomes

$$(\epsilon_{11} - \epsilon)^2 = \epsilon_{12}^2, \quad \epsilon_{11} - \epsilon = \pm \epsilon_{12} \quad (15)$$

and has the two roots

$$\epsilon = \epsilon_{11} \mp \epsilon_{12} \quad (16)$$

In this way we have obtained the second approximation to the proper value. The two-fold first approximation $E = E_1 + E_2$ of the proper value splits up into the two somewhat different values

$$E = E_1 + E_2 + (\epsilon_{11} \mp \epsilon_{12}) \frac{h^2}{8\pi^2 m} \quad (16a)$$

As we know ϵ we can calculate the ratio $\alpha : \beta$ from (12a). We get

$$\alpha \pm \beta = 0.$$

Inserted in (8) this gives

$$w = \alpha(u \mp v).$$

α is determined by the condition that w is to be normalised to 1. From this we get at once that $1 = 2\alpha^2$, hence

$$w = \frac{1}{\sqrt{2}}(u \mp v) \quad (17)$$

These are the two solutions of the first approximation with which the perturbation links up continuously. As we see, the favoured rôle of an "inner electron" which we had ascribed to one or other electron in u or v becomes lost in the combinations w which are alone of importance in the sequel: w is constructed symmetrically or anti-symmetrically in both electrons. Heisenberg expresses this by speaking of an "exchange of position" of the inner and outer electron. We wish to emphasise here that this exchange of position (*Platzwechsel* or *Austausch*), corresponding to the derivation of eqn. (17), is a direct analytical consequence of Schrödinger's theory of perturbation and does not signify a new physical hypothesis. The fact that the individual electrons cannot be numbered and distinguished individually in wave-mechanics (just as little as the light-waves of optics) is contained in the form of the wave-equation and forms the fundamental principle of the statistics of wave-mechanics, a question into which we cannot, however, enter here.

We shall now show that the two systems of states that are represented to a first approximation by (17) cannot combine with each other. To prove this we must link up with § 5, Chap. I, p. 46, by forming the matrix-element of the co-ordinate q . If q_{nm} vanishes for every component of q , this means that the transition $m \rightarrow n$ is forbidden (yields no radiation).

This leads us to test the transition from the state $\frac{1}{\sqrt{2}}(u + v)_m$ to the state $\frac{1}{\sqrt{2}}(u - v)_n$. The associated matrix element is

$$q_{nm} = \frac{1}{2} \iint (q_1 + q_2)(u + v)_m(u - v)_n d\tau_1 d\tau_2 \quad (18)$$

The integration is to be performed over the whole co-ordinate space of the q_1, q_2 's, with the space-elements $d\tau_1, d\tau_2$. To explain the factor $q_1 + q_2$ we recall that the matrix-element of the co-ordinates was originally calculated as an electrical moment (cf. p. 45), and that, for example, the x -co-ordinates of *both* electrons clearly contribute to the moment in the direction of the x -axis in such a way that the x -moment is to be formed with the "leverage" $x_1 + x_2$. Corresponding to this we have in (18) the factor $q_1 + q_2$ (but written in a somewhat more general form).

Instead of q_{nm} we first consider the integral

$$X = \int_{-a}^{+a} dx_1 \int_{-a}^{+a} dx_2 f(x_1, x_2) \quad (18a)$$

with the assumption

$$f(x_1, x_2) = -f(x_2, x_1) \quad (18b)$$

We argue as follows: the numerical value of X cannot change if we exchange the names of the variables of integration. So we must also have

$$X = \int_{-a}^{+a} dx_2 \int_{-a}^{+a} dx_1 f(x_2, x_1).$$

If we reverse the order of sequence of the integrations and take into account (18b), we get

$$X = - \int_{-a}^{+a} dx_1 \int_{-a}^{+a} dx_2 f(x_1, x_2).$$

Comparison with (18a) shows us directly that $X = 0$.

Precisely the same argument is applicable to (18), since $u - v$ and hence also $f(q_1, q_2) = (q_1 + q_2)(u + v)(u - v)$ changes sign if we exchange electron 1 for electron 2. Also the limits of integration in (18) are the same for both sets of three variables, being $\pm \infty$, just as in (18a) they were the same for both variables x_1, x_2 , namely $\pm a$. Consequently $q_{nm} = 0$ for every co-ordinate $q = x, y, z$ and for all values n, m .

The decree which forbids combinations between the states $u \mp v$ (*Interkombinations-Verbot*) itself points to the existence of the two term-systems of *ortho*- and *para-helium* which do not combine with each other. The important feature, however, is not that this decree does hold to a first approximation, as has just been proved, but that it holds *exactly*. This follows from the fact that by adding a second-approximation, that is, the function ϕ from eqn. (9) and all further approximations, the even or odd character, respectively, of the first approximation remains preserved with respect to the electrons 1 and 2. Here and in the whole of this section

we have taken no account of electron spin. In the case of He the spin of the electrons leads to an exceedingly weak combination between the two term-systems; this combination becomes increasingly stronger for the heavier atoms.

We now enquire into the difference in the position of the energy-levels that belong to the systems $u \mp v$. These energy-levels are given to a second approximation by (16a); so their difference is

$$\Delta E = E_{u-v} - E_{u+v} = -2\epsilon_{12} \cdot \frac{h^2}{8\pi^2 m} \quad (19)$$

On the other hand, the mean energy-level of these two comes out by (16a) as

$$E_m = \frac{1}{2}(E_{u-v} + E_{u+v}) = E_1 + E_2 + \epsilon_{11} \cdot \frac{h^2}{8\pi^2 m} \quad (20)$$

Eqn. (19) already allows us to come to a provisional decision as to how the allocation of $u \mp v$ is to be made to the actually observed ortho- and para-states. By (13) ϵ_{12} is positive, since s is positive throughout and u and v , on account of their normalisation, are at any rate positive on the average (for a more exact proof see C). Hence, by (19), ΔE becomes *negative*. From I, Fig. 84, p. 323, we see, however, that the ortho-levels always lie a little lower than the corresponding para-levels (we shall presently discuss the ground-level $n = 1$ in greater detail). The difference ortho minus para is therefore negative. Hence the states $u - v$ must correspond to the ortho-terms, the states $u + v$ to the para-terms. We arrive at the same result by considering the ground-state 1 S. We know from experiment that this is present only in the para-system. In our theoretical representation the ground-state is characterised by $n = 1$, that is, $u = v$. For this state, then, $u - v = 0$ and $u + v = 2u$. Thus the system given by $u - v$ has no state 1 S (a vanishing proper function never counts as a proper function). So, in agreement with our previous result, this system must be the *ortho-system*. On the other hand, the system belonging to $u + v$ has the ground-state u if the normalisation is appropriate; hence this is the *para-system*. (The normalising factor which in (17) was found to be $1/\sqrt{2}$ generally is clearly $1/2$ in the para-ground-state.) In the following section we shall show in yet a third way that our allocation of $u \mp v$ to the ortho- and para-systems is correct.

B. The Fine Structure of the Helium Spectrum

In the case of helium as in that of all the heavier elements the fine structure is essentially due to the "electron-spin." As has already been shown in Fig. 10 of § 9, Chap. I, this spin also contributes together with the relativity correction to the fine structure of the hydrogen spectrum. Dirac has shown (cf. § 10 of the present chapter) that the electron spin has *au fond* the same root as the relativity correction.

Whereas there was a mutual action of electrical origin $E_{12} = e^2/r_{12}$ between the centres of gravity of the two electrons, let there be a mutual action $M_{12} = M(\sigma_1, \sigma_2)$ of magnetic origin between the two electronic axes σ_1, σ_2 . Further, we imagine an external magnetic field of axis \mathbf{H} along which the axes σ_1 and σ_2 , each taken separately, set themselves. Let $\psi(\sigma)$ be the proper function of the individual electron in the field \mathbf{H} ; we imagine $\psi(\sigma)$ formed, for example, after the manner of § 3, Chap. I, as the proper function of the rotator (in space or in a plane). In it there occurs the magnetic quantum number which we shall here call m_1 or m_2 respectively. From Goudsmit and Uhlenbeck's hypothesis (Chap. I, p. 86), subject to the proof in § 10, we assume that m_1 and m_2 can have only the values $\pm \frac{1}{2}$. We then have the following combinations of quantum-numbers or proper functions, respectively.

m_1	m_2	m_s	Proper Functions.
$\frac{1}{2}$	$\frac{1}{2}$	1	$a = \psi(\frac{1}{2})\psi(\frac{1}{2})$
$\frac{1}{2}$	$-\frac{1}{2}$	0	$b = \psi(\frac{1}{2})\psi(-\frac{1}{2})$
$-\frac{1}{2}$	$\frac{1}{2}$	0	$c = \psi(-\frac{1}{2})\psi(\frac{1}{2})$
$-\frac{1}{2}$	$-\frac{1}{2}$	-1	$d = \psi(-\frac{1}{2})\psi(-\frac{1}{2})$

m_s is the quantum sum $m_1 + m_2$. The proper functions a, b, c, d are calculated to a first approximation as products of the partial functions, just as u and v in section A. Of these four proper functions a and d are not degenerate, just as earlier the ground state of helium was not degenerate. On the other hand, b and c have the same exchange-degeneracy as the excited states of helium, since their magnetic energies to a first approximation in the field \mathbf{H} , namely $(\mathbf{H}\sigma_1) + (\mathbf{H}\sigma_2)$, are themselves equal. To overcome this degeneracy we have to carry out a perturbation-calculation with the mutual action M_{12} analogous to that previously carried out for the mutual action E_{12} . In the determinative equation for the change of energy ϵ to a second approximation the same coefficients ϵ_{ik} occur as in (12a), except that now, of course, the perturbation quantity s is proportional to M_{12} . On account of its symmetry with respect to σ_1 and σ_2 we here also have $\epsilon_{22} = \epsilon_{11}$ and $\epsilon = \epsilon_{11} \mp \epsilon_{12}$. Hence the two combinations with which the proper functions of the second approximation are continuously connected are $b \mp c$. The proper functions of the second approximation that develop from the non-degenerate states a and d , on the other hand, are clearly continuously connected with a and d themselves. So we have three proper functions symmetrical in σ_1, σ_2 and one unsymmetrical, namely, written to a first approximation:

$$a, b + c, d \quad \text{and} \quad b - c, \text{ respectively.}$$

The first three form a triplet, the last a singlet of possibilities of orientation. The magnetic quantum-numbers are, by the preceding table,

$$m_s = +1, 0, -1 \quad \text{and} \quad 0, \text{ respectively.}$$

The proper functions of the whole problem, motion of the centre of gravity + orientation, are composed, to a first approximation, of the proper functions of the two partial problems. Four combinations are here possible, which we divide into the two groups (1) and (2) :

$$(1) \quad (u + v) \begin{pmatrix} a \\ b + c \\ d \end{pmatrix}, \quad (u - v)(b - c),$$

$$(2) \quad (u + v)(b - c), \quad (u - v) \begin{pmatrix} a \\ b + c \\ d \end{pmatrix}.$$

Our division into (1) and (2) is performed in such a way that all states of group (1) retain their sign when the two electrons are exchanged (symmetrical states), whereas those in group (2) change their signs (unsymmetrical states). In itself, that is, without a knowledge of the helium spectrum, the group (1) of states would be just as possible as the group (2). In eqn. (18), however, we have seen that combinations between symmetrical and unsymmetrical states are forbidden. Hence if a state of the group (1) were realised, it would never be able to combine with a state of the group (2).

The question as to which of the two groups occur in nature in the case of helium must, at the present state of the theory, be answered by experiment. In the first group the para-system, which we had to allocate to the proper functions $u + v$, would be a *triplet system*, the ortho-system a *singlet system*. As we know, the reverse is the case in reality. The combination (1) therefore drops out and the combination (2) is alone admissible empirically.

A deeper reason, namely the Pauli Exclusion Principle (§ 8, Chap. I, p. 85), becomes added to this. For in the ground state both electrons have, on account of $u = v$, exactly the same quantum numbers for the motion of the centre of gravity; hitherto we have had to mention only the principal quantum number n , but it is self-evident that the angular quantum numbers l and m must also have coincident values for both electrons; for from $n = 1$ it follows for both electrons that $l = 0$ and $m = 0$. By Pauli's principle the fourth quantum number m_s , which refers to the electronic axis, must be *different* for the two electrons. In our table this quantum number was called m_1 and m_2 . Hence in the ground state we cannot have $m_1 = m_2$; the orientations a ($m_1 = m_2 = +\frac{1}{2}$) and d ($m_1 = m_2 = -\frac{1}{2}$) are not allowed to occur in the ground state. Hence the ground state is not allowed to occur in the triplet system, as was the case in the combination (1). In the combination (2) the ground state of the triplet system drops out of itself on account of $u = v$, and the ground state of the singlet system is associated only with the orientation $b - c$ not with a and d .

We are still left with the problem of making assertions about the

magnitude and interval-relations of the fine structure. Heisenberg solved this on the basis of calculations supported on the old ideas of quantum orbits. The calculations are rather cumbersome and no longer correspond to the present state of the theory, which has been attained by Dirac's work on the spinning electron. We therefore give only the result of these calculations in the form of the following figures taken from Heisenberg (Fig. 28). All three figures represent the term $2p$ in the triplet system, with the three subsidiary terms p_2, p_1, p_0 and the relative quantum weight $2j + 1$ which are in the proportion $5:3:1$ (in Heisenberg's diagram the negative energy steps are drawn and are so in the reverse order to our term-steps). Let us first consider the limiting case $Z \rightarrow \infty$. Here we have the regular sequence of terms p_2, p_1, p_0 and the normal interval-ratio $2:1$ (cf. *Atombau*, p. 659); the term p_2 with the greatest weight 5

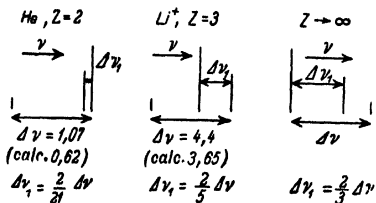


FIG. 28.—Triplet resolutions for He, Li^+ and very great nuclear charge. The weights of the levels are in the proportion $5:3:1$, the ratios of the intensity in the case of these strongly distorted triplets of He and Li^+ will possibly not have the "normal" values of the summation rules.

has the smallest numerical value. For Li^+ we have a "partially inverted" (*partiellverkehrte*) order of sequence; the "most intense" level p_2 lies in the middle. In the case of He the order of sequence is "inverted" throughout compared with the normal order; this triplet, which is highly degenerate in its interval-ratios was, as we know,

formerly regarded as a doublet experimentally but at the same time its apparent doublet character doubted on theoretical grounds (cf. *Atombau*, p. 599). The $\Delta\nu$'s marked under the figures denote cm^{-1} and are results of observation, below them and in brackets are the values calculated by Heisenberg. The data concerning $\Delta\nu_1$, that is, the distance p_2p_1 also follow from the formula developed by Heisenberg.

W. V. Houston* has applied these results to other atoms with two outer electrons, namely to Cd, Hg, Sn, Pb. The anomalous Zeeman effects of these atoms also allowed themselves to be calculated theoretically.

C. Quantitative Completion of the Perturbation Calculations

Following the general plan of the theory of perturbations we have so far developed, we have built up the first approximation from the proper functions for hydrogen ψ_1, ψ_n for the case of a Z -fold charged nucleus in eqns. (5) and (7). But this is unfavourable for the convergence of the perturbation calculation. The inner electron is indeed subject to the action of a Z -fold charged nucleus, but for the outer electron the nuclear charge is screened down to the value $Z - 1$ by the inner electron.

* Phys. Rev., 33, 297 (1927).

Heisenberg takes this into account by adding to the potential energy $V = -Ze^2/r$, in the case of each of the two electrons an additional energy $f(r)$ which has the value $+e^2/r$ in the region of the "orbit of the outer electron" but is constant in the region of the "orbit of the inner electron," and so corresponds to no force. Let the boundary between the two regions be $r = r_0$; the graph of $f(r)$ is given in Fig. 29. In view of this additional energy term we rewrite the differential equations (2) and

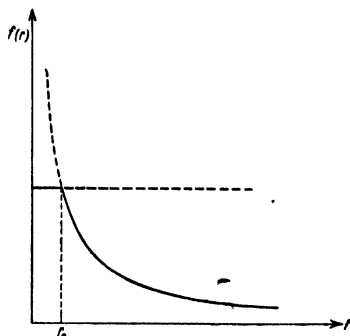


FIG. 29.—Diagrammatic representation of the perturbation function for the two helium electrons. $r < r_0$ denotes penetration into the "shell" of the "inner electron."

(3) of the first approximation. We first set $\lambda = 0$ in them. For since λ is a separation-parameter it would have to be of different signs in the two equations, but on account of the possibility of exchanging the two electrons it would simultaneously have to be of the same sign. Hence it necessarily follows that $\lambda = 0$. The two eqns. (2) and (3) are then, if we take $r < r_0$ in the first and $r > r_0$ in the second,

$$\Delta_1 \psi(1) + \frac{8\pi^2 m}{h^2} \left(E_1 - \frac{e^2}{r_0} + \frac{Ze^2}{r_1} \right) \psi(1) = 0 \quad (2')$$

$$\Delta_2 \psi(2) + \frac{8\pi^2 m}{h^2} \left(E_2 + \frac{(Z-1)e^2}{r_2} \right) \psi(2) = 0 \quad (3')$$

The n^{th} proper value of (3') now becomes

$$E_2 = - \frac{R h (Z-1)^2}{n^2} \quad (4')$$

The first proper value of (2') is

$$E_1 = - \frac{R h Z^2}{1^2} + \frac{e^2}{r_0} \quad (4'')$$

This fulfils the purpose aimed at above of making the electron near the nucleus subject to a Z -fold charge, that further away to a $(Z-1)$ -fold charge. The constant addition term e^2/r_0 in the case of the electron nearer the nucleus corresponds to an "external screening" (cf. *Atombau*, p. 544). Nevertheless, we cannot, strictly speaking, con-

sider the proper functions of the two electrons as having been calculated from eqns. (2') and (3'), which are meant only for $r < r_0$ and $r > r_0$, respectively, but we must consider them as coming from the equation

$$\Delta_i \psi(i) + \frac{8\pi^2 m}{h^2} \left(E_i - f(r_i) + \frac{Ze^2}{r_i} \right) \psi(i) = 0 \quad (5')$$

which is common to the two electrons ($i = 1$ or 2), f signifying the *same* function as that illustrated in Fig. 29; for otherwise we should introduce an asymmetry between the two electrons, which would destroy their interchangeability. By replacing these eqns. (5') by (2'); (3') for the purpose of calculating the proper values of a first approximation (4'), (4'') we have introduced an inaccuracy which, however, corrects itself in the next approximation. For the combination $w = \alpha u + \beta v$ formed from the products

$$u = \psi_1(1)\psi_n(2) \quad \text{or} \quad v = \psi_n(1)\psi_1(2),$$

we get the following differential equation (5') for the first approximation:

$$\left. \begin{aligned} L(w) \equiv \Delta w + \frac{8\pi^2 m}{h^2} \left(E_1 + E_2 - f(r_1) - f(r_2) \right) \\ + \frac{Ze^2}{r_1} + \frac{Ze^2}{r_2} \Big) w = 0 \end{aligned} \right\} \quad (6')$$

If, as a second approximation we set $\psi = w + \phi$, $E = E_1 + E_2 + \eta$, the differential equation for ϕ now runs

$$L(\phi) = (s - \epsilon)w, \quad (10')$$

as a comparison of (1) and (6') shows. The abbreviations s and ϵ are here defined by

$$s = C \left(\frac{e^2}{r_{12}} - f(r_1) - f(r_2) \right), \quad \epsilon = C\eta, \quad C = \frac{8\pi^2 m}{h^2} \quad (11')$$

instead of by (11); the products $\epsilon\phi$ and $s\phi$, being small quantities of higher order have been omitted. Further, the meaning of the quantities $\epsilon_{11} = \epsilon_{22}$ and ϵ_{12} from eqn. (13) must be changed. We now pass on to consider the calculation of

$$\eta_{12} = \frac{\epsilon_{12}}{C} \quad (13')$$

a little more closely.

We form the functions ψ_1, ψ_n that occur in u and v approximately as solutions of eqns. (2') and (3'), that is, as hydrogen proper-functions (ψ_1 with the nuclear charge Z , ψ_n with the nuclear charge $Z - 1$) and, taking up the normalising factors into the radial-components R_1, R_n , we set

$$\psi_1 = R_1 \cdot 1, \quad \psi_n = R_n P_l^m(\cos \theta) e^{im\phi}, \\ |m| \leq l, \quad l \leq n - 1.$$

We form η_{12} according to the scheme of ϵ_{12} in eqn. (13) and perform the integration first with respect to θ_1, ϕ_1 , then with respect to θ_2, ϕ_2 , and finally with respect to r_1, r_2 , in each case writing down only such factors

as contain the corresponding variables of integration. The first step gives

$$\int_0^\pi \int_0^{2\pi} s \cdot 1 \cdot P_l^m(\cos \theta_1) e^{-im\phi_1} \sin \theta_1 d\theta_1 d\phi_1 \quad (21)$$

The factor 1 is due to $\psi_1(1)$ in u , the two following factors are due to $\psi_n^*(1)$ in v^* .† In the integration the terms $f(r_1)$ and $f(r_2)$ which do not contain θ_1, ϕ_1 drop out† in s . With

$$\begin{aligned} r_{12}^2 &= r_1^2 + r_2^2 - 2r_1 r_2 \cos \Theta, \\ \cos \Theta &= \cos \theta_1 \cos \theta_2 + \sin \theta_1 \sin \theta_2 \cos(\phi_1 - \phi_2), \\ \frac{1}{r_{12}} &= \frac{1}{[r]} \sum_0^\infty x^v P_v(\cos \Theta), \end{aligned}$$

where $[r]$ denotes the greater of the two radii r_1, r_2 and x the ratio of the smaller to the greater r , we get from (21)

$$\frac{e^2}{[r]} \sum_0^\infty x^v \int_0^\pi \int_0^{2\pi} P_v(\cos \Theta) P_l^m(\cos \theta_1) e^{-im\phi_1} \sin \theta_1 d\theta_1 d\phi_1 \quad (22)$$

For $P_v(\cos \Theta)$ we use the addition theorem of spherical harmonics, eqn. (2) of § 8, Chap I:

$$P_v(\cos \Theta) = \sum_{-\nu}^{+\nu} \frac{(\nu - |\mu|)!}{(\nu + |\mu|)!} P_\nu^\mu(\cos \theta_1) P_\nu^\mu(\cos \theta_2) e^{i\mu(\phi_1 - \phi_2)}.$$

Of all the terms of this series there remains, on account of the integration with respect to ϕ_1 , only the term $\mu = m$, so that (22) becomes

$$2\pi \frac{e^2}{[r]} \sum_0^\infty x^v \frac{(\nu - m)!}{(\nu + m)!} P_\nu^m(\cos \theta_2) e^{-im\phi_2} \cdot \Pi \quad (23)$$

where
$$\Pi = \int_0^\pi P_\nu^m(\cos \theta_1) P_l^m \cos(\theta_1) \sin \theta_1 d\theta_1.$$

But by eqns. (24) and (30) of § 6, Chap. I, we have

$$\Pi = \begin{cases} 0 & \dots \nu \neq l, \\ \frac{2}{2l+1} \cdot \frac{(l+m)!}{(l-m)!} & \dots \nu = l. \end{cases}$$

† It is to be noticed that when the proper functions are written in the complex form the conditions of orthogonality (11a) are also to be written in the complex form, for example

$$\int (s - e) w u^* d\tau = 0.$$

Accordingly the definition of e_{lk} in (18) must also be modified, for example,

$$e_{12} = \int s u v^* d\tau, \quad e_{13} = \int s u^* v d\tau = e_{13}^*.$$

† We exclude the case $l = 0$ (S-term).

Consequently (23) reduces to a single member (the factorial factor again cancels out)

$$4\pi \frac{e^2}{[r]} \frac{x^l}{2l+1} P_l^m(\cos \theta_2) e^{-im\phi_2} \quad (24)$$

The second step consists in multiplying this expression by

$$P_l^m(\cos \theta_2) e^{im\phi_2} \cdot 1,$$

where the 1 arises from v^* and the preceding factors from u . In the integration with respect to θ_2, ϕ_2 (24) becomes

$$8\pi^2 \frac{e^2}{[r]} \frac{x^l}{2l+1} \Pi,$$

with the value of Π just now given, which does not vanish. Thus the second step gives

$$16\pi^2 \frac{e^2}{[r]} \frac{x^l}{(2l+1)^2} \frac{(l+m)!}{(l-m)!}$$

We now come to the third step, that is, to the integration with respect to r_1, r_2 . The expression to be integrated is [cf. (13')]

$$\eta_{12} = \left. \begin{aligned} & \frac{16\pi^2 e^2}{(2l+1)^2} \frac{(l+m)!}{(l-m)!} \int_0^\infty \frac{x^l}{[r]} \cdot R(r_1, Z) R_n(r_1, Z-1) r_1^2 dr_1 \\ & \cdot \int_0^\infty R_1(r_2, Z) \cdot R_n(r_2, Z-1) r_2^2 dr_2 \end{aligned} \right\} \quad (25)$$

Here we have (a = radius of the first hydrogen orbit)

$$R_1(r, Z) = N_1(Z) e^{-\rho/2} L_1^1(\rho) \dots \rho = \frac{2Z}{a} r,$$

$$R_n(r, Z-1) = N_n(Z-1) e^{-\rho'/2} L_{n+l}^{2l+1}(\rho') \dots \rho' = \frac{2(Z-1)}{na} r.$$

The last formulæ hold for $r = r_1$ as well as for $r = r_2$. The factors N contain both the normalisation for the dependence on θ, ϕ as well as on r and run

$$\begin{aligned} N_1^2(Z) &= \left(\frac{2Z}{a}\right)^3 \cdot \frac{1}{2} \cdot \frac{1}{2\pi} \cdot \frac{1}{2}, \\ N_n^2(Z-1) &= \left(\frac{2(Z-1)}{na}\right)^3 \frac{(n-l-1)!}{2n(n+l)!^3} \cdot \frac{1}{2\pi} \cdot \frac{2l+1}{2} \cdot \frac{(l-m)!}{(l+m)!} \end{aligned}$$

We restrict ourselves to the case $n = l+1$ (circular orbits in the earlier sense), where the polynomials L become of zero degree, since then

$$L_{n+l}^{2l+1} = L_{2l+1}^{2l+1} = \text{const.} = -(2l+1)!$$

holds. According to the definition of the quantities x and $[r]$ the integrals in (25) are to be performed as follows:

$$\int_0^\infty r_1 e^{-\sigma r_1} dr_1 \left\{ \int_0^{r_1} r_2^{2l+2} e^{-\sigma r_2} dr_2 + r_1^{2l+1} \int_{r_1}^\infty r_2 e^{-\sigma r_2} dr_2 \right\} \quad (26)$$

Here α denotes the abbreviation

$$\alpha = \frac{1}{a} \left(Z + \frac{Z-1}{n} \right);$$

we must imagine the constant factor

$$16\pi^2 e^2 \frac{(l+m)!}{(l-m)!} [2l]!^2 N_1^2(Z) N_n^2(Z-1) \cdot \left(\frac{2(Z-1)}{na} \right)^{2l}$$

to be added. The rather laborious working out of (26) gives

$$\alpha^{-2l-5} \cdot 2^{-2l-5} \cdot (2l+2)! \cdot (2l+5).$$

Taking into account the constant factor we get

$$\eta_{12} = 8R\hbar \frac{Z^3(Z-1)^{2l+3}}{[(l+2)Z-1]^{2l+3}} \frac{(l+1)^2(2l+5)}{2l+1} \quad (27)$$

where R = Rydberg's number.

This is our final result; $2\eta_{12}$ gives to the approximation in question the energy-difference between para- and ortho-levels [cf. (19)], with the assumption $n = l + 1$ for helium-like atoms of nuclear charge Z . Numerical values of $\eta_{12}/R\hbar$ have been calculated by Heisenberg. We have tabulated them in Table 5; the bracketed values are the corresponding values obtained by observation.

TABLE 5.

	$l = 1$ (2P).	$l = 2$ (3D).	$l = 3$ (4F).
$Z = 2$	$7.65 \cdot 10^{-3}$ ($9.4 \cdot 10^{-3}$)	$2.57 \cdot 10^{-5}$ ($1.84 \cdot 10^{-5}$)	$5.25 \cdot 10^{-8}$ —
$Z = 3$	$3.07 \cdot 10^{-3}$ —	$1.89 \cdot 10^{-4}$ ($1.34 \cdot 10^{-4}$)	$6.95 \cdot 10^{-7}$ —

There is agreement between the results of observation and calculation in order of magnitude, particularly for greater values of l ; we have reason for believing that this agreement can be made perfect by carrying the perturbation calculation still further. More important than this quantitative agreement is the qualitative agreement of the structure and fine-structure derived in sections A and B with the results of experiment, and the numerical calculation of the ground term (the ionisation potential, cf. the introduction to this §). We may therefore say that the helium problem has been solved by wave-mechanics and that therefore (cf. § 1, p. 144) wave-mechanics need no longer be daunted by the problem of three bodies.

The method that Heisenberg has developed for the helium problem may be applied to the problem of several bodies, that is, to atoms with more than two electrons. With the increase in the number of electrons the exchange-degeneracy (*Austausch Entartung*) increases; the determinant for calculating the energy-perturbation ϵ becomes of degree $N!$

for N electrons; in the general case several multiplet systems occur instead of a singlet or a triplet system. The detailed investigation of these systems by the principles of the theory of groups has been advanced particularly by the papers of W. Heisenberg, E. Wigner, J. von Neumann, and H. Heitler. * In the general case, too, the possible proper functions resolve into several systems unable to combine with one another, which correspond to the groups (1) and (2) of p. 233. In nature only one of these groups, namely that of anti-symmetrical character, is actually realised. This anti-symmetry of the proper functions is the wave-mechanical equivalent to Pauli's principle and at the same time the most general way of formulating it.

The same methods have led Heitler and London * to a solution of the two-electron problem of the H_2 -molecule, namely, to a determination of the energy and the moment of inertia of these molecules in the ground state. Finally, it can be predicted by theory that on account of the spin of the two hydrogen nuclei this ground-state can occur in two different forms, an ortho- and a para- form, according as the spins are parallel or anti-parallel. The ratio of these two forms of H_2 , which is 3:1 at higher temperatures, may be altered experimentally at low temperatures, as has been shown by Bonhöffer and Harteck † and also by Eucken. ‡

§ 2. Wave-Mechanical Transcription of Classical Quantities

We start from Schrödinger's definition of density $\rho = uu^*$ and form with it the wave-mechanical mean value of the co-ordinate q or, as we may say, the co-ordinate of the centre of gravity of the wave-picture given by u :

$$\bar{q} = \int q \rho d\tau = \int quu^* d\tau \quad . \quad . \quad . \quad (1)$$

For the sake of brevity we make our calculations with an individual point-mass and so take $d\tau$ as standing for the three-dimensional element of volume; we remark, however, that all the following calculations may easily be extended to arbitrary systems of point-masses or electrons.

If in (1) we make u and u^* in particular equal to the n^{th} proper function of a wave-mechanical problem defined in some way, then \bar{q} simultaneously denotes the n^{th} diagonal-element q_{nn} of the co-ordinate-matrix. We extend (1) by replacing u by u_n and u^* by u_m^* and so obtain the general matrix element

$$q_{nm} = \int qu_n u_m^* d\tau \quad . \quad . \quad . \quad (1a)$$

* Zeitschr. f. Phys., 44, 455 (1927).

† Sitzungsberichte der Preussischen Akademie, Berlin, p. 108 (1929).

‡ Naturwiss., 17, 182 (1929).

§ 9. Wave-Mechanical Transcription of Classical Quantities 241

A. Energy and Momentum

We define the wave-mechanical mean value of the potential energy V analogously by *

$$U = \int V u u^* d\tau \quad . \quad . \quad . \quad . \quad . \quad (2)$$

If in particular $u = u_n$ (n^{th} proper function), then $U = V_{nn}$ (diagonal element of the V -matrix). Here, too, we may generalise and pass on to the arbitrary matrix-element V_{nm} .

If V is independent of t the energy-law in ordinary mechanics holds in the simple form

$$E = K + V \quad (K = \text{kinetic energy}) \quad . \quad . \quad . \quad (3)$$

and in wave-mechanics the differential equation for u reduces to the wave-equation, which does not involve the time,

$$\frac{\hbar^2}{8\pi^2 m} \Delta \psi + (E - V) \psi = 0.$$

By multiplying with ψ^* and integrating we get from this

$$\frac{\hbar^2}{8\pi^2 m} \int \Delta \psi \psi^* d\tau + E - U = 0 \quad . \quad . \quad . \quad (4)$$

We have here taken into account that owing to the constancy of E and of the normalising condition for ψ , we get

$$\int E \psi \psi^* d\tau = E,$$

and that because of

$$u = \psi e^{\frac{2\pi i}{\hbar} Et}, \quad u^* = \psi^* e^{-\frac{2\pi i}{\hbar} Et},$$

the mean value U defined in (2) is now identical with

$$U = \int V \psi \psi^* d\tau.$$

A comparison of (4) with (3) invites us to regard as the wave-mechanical analogy to the kinetic energy the expression

$$K = \frac{1}{2m} \left(\frac{\hbar}{2\pi i} \right)^2 \int \Delta \psi \psi^* d\tau \quad . \quad . \quad . \quad (5)$$

To bring this expression externally into correspondence with the elementarily definition

$$K = \frac{1}{2m} (p_x^2 + p_y^2 + p_z^2) \quad . \quad . \quad . \quad (5a)$$

we could re-write (5) by Green's theorem in the form

$$K = \frac{1}{2m} \left(\frac{\hbar}{2\pi i} \right)^2 \int \left(\frac{\partial \psi}{\partial x} \frac{\partial \psi^*}{\partial x} + \frac{\partial \psi}{\partial y} \frac{\partial \psi^*}{\partial y} + \frac{\partial \psi}{\partial z} \frac{\partial \psi^*}{\partial z} \right) d\tau \quad . \quad (5b)$$

But we shall soon learn that the true analogy to (5a) occurs not in the form (5b) but actually in that of (5).

* The symbol U here introduced has nothing in common, of course, with the earlier symbol $V = E_0 + U$.

First we must find the wave-mechanical expression analogous to the momentum p 89. Its definition is really already contained in that of the current S , eqn. (8), § 8, Chap. I. For if we multiply S by m/e we pass from the wave-mechanical or probable value of the transported charge to the correspondingly defined value of the transported mass. This is the *specific momentum* for a definite point of the wave-mechanical field, just as S was the specific current. Thus we obtain

$$\frac{m}{e}S = \frac{\hbar}{4\pi i}(u \operatorname{grad} u^* - u^* \operatorname{grad} u) \quad (6)$$

To obtain the *total momentum* from this we have only to integrate with respect to $d\tau$,

$$p = \frac{\hbar}{4\pi i} \int (u \operatorname{grad} u^* - u^* \operatorname{grad} u) d\tau \quad (6a)$$

At the same time we obviously have

$$0 = \frac{\hbar}{4\pi i} \int (u \operatorname{grad} u^* + u^* \operatorname{grad} u) d\tau \quad (6b)$$

for the integral in this last equation may be written in the form

$$\int \operatorname{grad} (uu^*) d\tau,$$

and may be reduced by integration by parts to a surface-integral at infinity which at any rate certainly vanishes for the solutions u defined with discrete* proper values. By adding the two eqns. (6a, b) together we then get as the simplified form of (6a)

$$p = \frac{\hbar}{2\pi i} \int u \operatorname{grad} u^* d\tau \quad (7)$$

This is our wave-mechanical paraphrase of the classical conception of momentum. We extend (7) to represent the momentum-matrix

$$p_{nm} = \frac{\hbar}{2\pi i} \int u_n \operatorname{grad} u_m^* d\tau \quad (7a)$$

We shall give (7) a symbolic form that will be important in the sequel:

$$\left. \begin{aligned} p &= \int u \pi u^* d\tau \\ \pi &= \frac{\hbar}{2\pi i} \frac{\partial}{\partial q} = \frac{\hbar}{2\pi i} \operatorname{grad} \end{aligned} \right\} \quad (8)$$

We call π the "momentum-operator." It is to be applied to the quantity u^* which follows it. The order of sequence in which the operator π and

*In the case of the proper functions of the continuous spectrum the order of vanishing at infinity is much less, so that a certain caution is advisable. In the sequel we assume discrete proper values in order that we may always omit the surface-integrals at infinity, as will be done in individual cases without further explanation.

§ 9. Wave-Mechanical Transcription of Classical Quantities 243

the factors u, u^* occur is, of course, of essential importance. Operators are not commutative. On the other hand, we can exchange the order of sequence in eqn. (1) in which only ordinary factors occur, and we can obviously bring this equation into the form

$$\bar{q} = \int u q u^* d\tau$$

which is analogous to (8). In Dirac's terminology ordinary numbers are called "c-numbers" (classical numbers), while operators are called "q-numbers" (numbers arising from the quantum theory).

The introduction of the operator π is connected with the association of the momentum p with the action-function S in classical theory [S will further be used in a sense different from that in eqn. (6)], namely

$$p = \frac{\partial S}{\partial q}$$

We recall the original introduction of the wave-equation in § 1, Chap. I, and its derivation from the Hamiltonian equation

$$\sum p^2 = \sum \left(\frac{\partial S}{\partial q} \right)^2 = 2m(E - V).$$

If we here insert the symbol π , that is, the operator $(\hbar/2\pi i) \partial/\partial q$ for p or the differential quotient $\partial S/\partial q$ respectively, and apply the whole equation to ψ , then we actually get the wave-equation in the exact form (11) of p. 5.

This at the same time justifies our expression (5) for the kinetic energy. For if in each case we insert π in place of p in the classical expression (5a) and integrate after the manner of eqn. (8), we get

$$\left. \begin{aligned} K &= \frac{1}{2m} \int u (\pi_x^2 + \pi_y^2 + \pi_z^2) u^* d\tau \\ &= \frac{1}{2m} \left(\frac{\hbar}{2\pi i} \right)^2 \int u \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right) u^* d\tau \end{aligned} \right\} \quad (8a)$$

and this is, in actual fact, exactly our expression (5), except that now u and u^* have taken the place of the earlier quantities ψ and ψ^* which were a little more special.

To illustrate our definition (7) of the wave-mechanical momentum we consider the case of an electron which moves under no forces in the x -direction. We represent it by a plane wave (cf. also § 1, Chap. I)

$$u = ae^{-ikx + i\omega t}$$

Here k and ω are defined by

$$\left. \begin{aligned} k &= \frac{2\pi}{\hbar} \sqrt{2mK}, \quad \omega = \frac{2\pi}{\hbar} E \\ E &= mc^2 + K = \text{rest energy} \\ &\quad + \text{kinetic energy of the electron} \end{aligned} \right\} \quad (8b)$$

On account of our special form of u we obtain for the density ρ and the current S

$$\rho = e \cdot a^2, \quad S_x = \frac{h}{2\pi} \frac{e}{m} a^2 k, \quad S_y = S_z = 0,$$

and for p , by (7)

$$p = p_x = \frac{h}{2\pi} a^2 k \int d\tau.$$

But $\int a^2 d\tau = 1$ (normalising condition),* hence

$$p_x = \frac{hk}{2\pi} = \frac{h}{\lambda} \quad . \quad . \quad . \quad . \quad (8c)$$

where λ , as in § 1, Chap. I, signifies the de Broglie wave-length of the particle. The fact that this quantity plays the part of the momentum is easily seen. For we have

$$\frac{1}{2m} p_x^2 = K$$

since by (8b)

$$\frac{1}{2m} \cdot \frac{h^2 k^2}{(2\pi)^2} = \frac{1}{2m} \left(\frac{h}{2\pi} \right)^2 \left(\frac{2\pi}{h} \right)^2 \cdot 2m \cdot K = K.$$

B. Law of Conservation of the Centre of Mass†

We split up the law of conservation of the centre of mass of classical mechanics into the two equations ("definition of momentum" and "law of momentum"):

$$p = m\dot{q} \text{ and } \dot{p} = -\text{grad } V.$$

We transcribe them wave-mechanically by taking \bar{q} and p to stand for the expression (1) and (7) and forming grad V according to the process of eqn. (2):

$$p = m \dot{\bar{q}} \quad . \quad . \quad . \quad (9) \quad \dot{p} = - \int u u^* \text{grad } V d\tau \quad . \quad (10)$$

To prove this we calculate from (1) and (7)

$$\dot{\bar{q}} = \int q(\dot{u}u^* + u\dot{u}^*)d\tau \quad . \quad . \quad . \quad . \quad (9a)$$

$$\dot{p} = \frac{h}{2\pi i} \int (i \text{grad } u^* + u \text{grad } \dot{u}^*)d\tau \quad . \quad . \quad (10a)$$

(observing that q in (1) and grad in (7) are independent of the time). Here we must insert the differential quotients with respect to the time

* Thus a becomes infinitely small, as is to be expected, for the density of the electron in a plane wave is spread evenly over the whole of infinite space (cf. also § 3, E, and the end of § 7, D, of Chap. I).

† Cf. the somewhat more special statement of this law for one dimension by P. Ehrenfest, *Zeitschr. f. Phys.*, **45**, 455 (1927); also A. E. Ruark, *Phys. Rev.*, **31**, 583 (1928).

from eqns. (13) and (13a) of § 5, Chap. I:

$$\left. \begin{aligned} \dot{u} &= \frac{\hbar}{4\pi i m} \left(\Delta u - \frac{8\pi^2 m}{\hbar^2} V u \right) \\ \dot{u}^* &= - \frac{\hbar}{4\pi i m} \left(\Delta u^* - \frac{8\pi^2 m}{\hbar^2} V u^* \right) \end{aligned} \right\} \quad (11)$$

Then the term in V cancels out in (9a) and there remains

$$m \dot{q} = \frac{\hbar}{4\pi i} \int q (u^* \Delta u - u \Delta u^*) d\tau \quad (12)$$

But

$$\begin{aligned} u^* \Delta u - u \Delta u^* &= u^* \operatorname{div} \operatorname{grad} u - u \operatorname{div} \operatorname{grad} u^* \\ &= \operatorname{div} (u^* \operatorname{grad} u - u \operatorname{grad} u^*). \end{aligned}$$

By integrating (12) by parts we do away with the "divergence." If, for example, $q = x$, then, of the three differential coefficients, of which div is composed in rectangular co-ordinates only that with respect to x gives a value different from zero. In this case we get from (12)

$$m \dot{x} = - \frac{\hbar}{4\pi i} \int \left(u^* \frac{\partial u}{\partial x} - u \frac{\partial u^*}{\partial x} \right) d\tau;$$

but by (6a) the right-hand side is nothing other than p_x . Hence eqn. (9) is proved.

To prove eqn. (10) we first form (10a) by integration by parts of the second term on the right-hand side. We obtain

$$\dot{p} = \frac{\hbar}{2\pi i} \int (u \operatorname{grad} u^* - u^* \operatorname{grad} u) d\tau.$$

Substituting from (11) we get

$$\dot{p} = - \frac{\hbar^2}{8\pi^2 m} (\text{I} + \text{II}) \quad (13)$$

$$\text{I} = \int (\Delta u \operatorname{grad} u^* + \Delta u^* \operatorname{grad} u) d\tau \quad (13a)$$

$$\text{II} = - \frac{8\pi^2 m}{\hbar^2} \int (V u \operatorname{grad} u^* + V u^* \operatorname{grad} u) d\tau \quad (13b)$$

We first occupy ourselves with II. By integrating again by parts, it follows that

$$\begin{aligned} \int V u \operatorname{grad} u^* d\tau &= - \int u^* \operatorname{grad} (V u) d\tau = - \int u^* u \operatorname{grad} V d\tau \\ &\quad - \int u^* V \operatorname{grad} u d\tau. \end{aligned}$$

If this is substituted in II the last term cancels out and there remains

$$\text{II} = \frac{8\pi^2 m}{\hbar^2} \int u^* u \operatorname{grad} V d\tau.$$

If we assume that $I = 0$, which we shall presently prove, (13) becomes

$$\dot{p} = - \int uu^* \text{grad } V d\tau$$

and this is precisely the eqn. (10) which was to be proved, that is, *the law of momentum or the second part of the law of the centre of mass.*

The proof that I vanishes rests on the following identity:

$$\Delta u \text{ grad } u^* + \Delta u^* \text{ grad } u = \text{div } T \quad (14)$$

Here T is a symmetrical tensor* with the components

$$T_{ik} = \frac{\partial u}{\partial x_i} \frac{\partial u^*}{\partial x_k} + \frac{\partial u^*}{\partial x_i} \frac{\partial u}{\partial x_k} - \delta_{ik} \sum_j \frac{\partial u}{\partial x_j} \frac{\partial u^*}{\partial x_j} \quad (15)$$

i, k, j have the numerical values 1, 2, 3, where we are to imagine $x_1 = x$, $x_2 = y$, $x_3 = z$. δ_{ik} is the well-known unit tensor; $\text{div } T$ signifies a vector whose i -component is given by

$$\sum_k \frac{\partial}{\partial x_k} T_{ik}$$

and hence, according to (15), by

$$\begin{aligned} \frac{\partial u}{\partial x_i} \sum_k \frac{\partial^2 u^*}{\partial x_k^2} + \frac{\partial u^*}{\partial x_i} \sum_k \frac{\partial^2 u}{\partial x_k^2} + \sum_k \frac{\partial^2 u}{\partial x_i \partial x_k} \frac{\partial u^*}{\partial x_k} + \sum_k \frac{\partial^2 u^*}{\partial x_i \partial x_k} \frac{\partial u}{\partial x_k} \\ - \frac{\partial}{\partial x_i} \sum_k \frac{\partial u}{\partial x_k} \frac{\partial u^*}{\partial x_k} \end{aligned}$$

In the last term, on account of the meaning of δ_{ik} , we have replaced

$$\delta_{ik} \sum_k \frac{\partial}{\partial x_k} \text{ by } \frac{\partial}{\partial x_i}$$

and, for convenience, have written k for the summation index j used in (15). If we carry out the differentiation in this last term we see at once that it cancels out with the two preceding terms. But the first two terms are identical with the i -component of the left-hand side of (14). This proves (14), and it also shows that I vanishes. For by (13a) and (14) we have

$$I = \int \text{div } T d\tau = \sum \int \frac{\partial}{\partial x_k} T_{ik} d\tau = 0.$$

C. The Law of Sectorial Areas (Kepler's Third Law)

To formulate this law we require the wave-mechanical transcription of the conception "moment of momentum" on the one hand and "moment of a force" on the other. The moment of momentum with respect to an arbitrary origin from which the radius vector r is reckoned,

* Our T is closely related to the tensor S introduced by Schrödinger, *Ann. d. Phys.*, **82**, 265 (1927).

§ 9. Wave-Mechanical Transcription of Classical Quantities 247

is classically defined by $[\mathbf{r}\mathbf{p}]$. By eqn. (8) it is symbolised wave-mechanically by

$$P = \int u[\mathbf{r}\mathbf{p}]u^*d\tau = \frac{h}{2\pi i} \int u[\mathbf{r} \text{ grad } u^*]d\tau \quad (16)$$

The moment of a force is given classically by $-\mathbf{r} \text{ grad } V$ with respect to the same origin, if a potential V exists, and translated into the language of wave-mechanics it is

$$M = - \int u[\mathbf{r} \text{ grad } V]u^*d\tau \quad (17)$$

Now, the law of areas asserts that

$$\dot{P} = M \quad (18)$$

To prove this we must work out

$$\dot{P} = \frac{h}{2\pi i} \left\{ \int \dot{u}[\mathbf{r} \text{ grad } u^*]d\tau + \int u[\mathbf{r} \text{ grad } \dot{u}^*]d\tau \right\} \quad (19)$$

The second integral is next transformed in accordance with the scheme

$$\int u x \frac{\partial v}{\partial y} d\tau = - \int \frac{\partial u}{\partial y} x v d\tau \quad (19a)$$

With $v = \dot{u}^*$, this gives

$$\int u[\mathbf{r} \text{ grad } \dot{u}^*]d\tau = - \int \dot{u}^*[\mathbf{r} \text{ grad } u]d\tau$$

Hence we get, instead of (19),

$$\dot{P} = \frac{h}{2\pi i} \left\{ \int \dot{u}[\mathbf{r} \text{ grad } u^*]d\tau - \int \dot{u}^*[\mathbf{r} \text{ grad } u]d\tau \right\} \quad (20)$$

We have now to enlist the help of eqns. (11). We write, in a similar way to (13),

$$\dot{P} = - \frac{h^2}{8\pi^2 m} \{I + II\} \quad (21)$$

$$I = \int \{ \Delta u[\mathbf{r} \text{ grad } u^*] + \Delta u^*[\mathbf{r} \text{ grad } u] \} d\tau \quad (21a)$$

$$II = - \frac{8\pi^2 m}{h^2} \int \{ V u[\mathbf{r} \text{ grad } u^*] + V u^*[\mathbf{r} \text{ grad } u] \} d\tau \quad (21b)$$

Here, too, the integral I vanishes. For, taking (14) into consideration, we have

$$I = \int [\mathbf{r}, \Delta u \text{ grad } u^* + \Delta u^* \text{ grad } u] d\tau = \int [\mathbf{r}, \text{div } \mathbf{T}] d\tau$$

By considering the component given by the indices ik of the vector-product and using the index notation of (15) we get

$$I = \int \left\{ x_i \sum_j \frac{\partial}{\partial x_j} T_{kj} - x_k \sum_j \frac{\partial}{\partial x_j} T_{ij} \right\} d\tau$$

From this it follows by the rule of (19a) that

$$I = - \int \{T_{ki} - T_{ik}\} d\tau \quad (22)$$

and this is in actual fact equal to zero on account of the symmetry of the tensor T.

We next apply the rule (19a) to the first integral in II :

$$\begin{aligned} \int V u [\mathbf{r} \text{ grad } u^*] d\tau &= - \int u^* [\mathbf{r} \text{ grad } (Vu)] d\tau \\ &= - \int u u^* [\mathbf{r} \text{ grad } V] d\tau - \int u^* V [\mathbf{r} \text{ grad } u] d\tau. \end{aligned}$$

Here the last integral cancels out with the second integral of (21b). Hence we get from (21b) and (17)

$$II = \frac{8\pi^2 m}{h^2} \int u u^* [\mathbf{r} \text{ grad } V] d\tau = - \frac{8\pi^2 m}{h^2} M \quad (23)$$

Through this result eqn. (21) becomes identical with (18).

Thus our *law of area* is proved in its general form. The special form, such as occurs in the Kepler problem as the law of conservation of moment of momentum or the law of the constancy of areal velocity, occurs if we assume V to be a function of r alone (central force). Then

$$\text{grad } V = \frac{r}{r} \frac{dV}{dr},$$

and hence

$$[\mathbf{r} \text{ grad } V] = \frac{1}{r} \frac{dV}{dr} [\mathbf{r} \mathbf{r}] = 0, \quad M = 0 \quad (24)$$

In particular it follows that

$$\dot{P} = 0, \quad P = \text{Const.} \quad (25)$$

D. Law of the Virial

Whereas the *moment of a force* is defined by the vector product of \mathbf{r} and $-\text{grad } V$, the *virial* is defined by the scalar product of \mathbf{r} and $-\text{grad } V$. Calling the virial R (say) we write classically

$$R = - (\mathbf{r} \text{ grad } V).$$

The wave-mechanical paraphrase will then be

$$R = - \iint u u^* (\mathbf{r} \text{ grad } V) d\tau \quad (26)$$

In classical mechanics the name *law of virial* is given, as we know, to the following relation, which follows directly from the law of momentum :

$$\frac{d}{dt}(\mathbf{r} \mathbf{p}) = 2K + R \quad (27)$$

Applied to periodic or quasi-periodic orbits it allows us to infer that the mean value of twice the kinetic energy is equal and opposite to the mean value of the virial. The same inference is also useful for wave-mechanics.

We prove the wave-mechanical analogue of eqn. (27) by starting from the quantity constructed analogously to (\mathbf{r}, p) [cf. (8)],

$$Q = \int u(\mathbf{r}) u^* d\tau = \frac{h}{2\pi i} \int u(\mathbf{r} \text{ grad } u^*) d\tau \quad (28)$$

Its derivative with respect to t corresponds to the left-hand side of (27):

$$\begin{aligned} \dot{Q} &= \frac{h}{2\pi i} \int \{ \dot{u}(\mathbf{r} \text{ grad } u^*) + u(\mathbf{r} \text{ grad } \dot{u}^*) \} d\tau \\ &= \frac{h}{2\pi i} \int \{ \dot{u}(\mathbf{r} \text{ grad } u^*) - \dot{u}^* [3u + (\mathbf{r} \text{ grad } u)] \} d\tau \end{aligned} \quad (29)$$

In the last transformation use has been made of the following relation which is related to (19a):

$$\begin{aligned} \int u \sum x_i \frac{\partial v}{\partial x_i} d\tau &= - \int v \sum \frac{\partial}{\partial x_i} (x_i u) d\tau \\ &= - \int v \left(3u + \sum x_i \frac{\partial u}{\partial x_i} \right) d\tau \dots \begin{cases} i = 1, 2, 3, \\ x_1 = x, x_2 = y, x_3 = z \end{cases} \end{aligned} \quad (29a)$$

If we insert \dot{u} and \dot{u}^* from (11) in (29) we get

$$\dot{Q} = - \frac{h^2}{8\pi^2 m} (I + II) \quad (30)$$

with the abbreviations

$$\begin{aligned} I &= \int \Delta u (\mathbf{r} \text{ grad } u^*) + \Delta u^* [3u + (\mathbf{r} \text{ grad } u)] d\tau \\ II &= - \frac{8\pi^2 m}{h^2} \int \{ \nabla u (\mathbf{r} \text{ grad } u^*) + \nabla u^* [3u + (\mathbf{r} \text{ grad } u)] \} d\tau \end{aligned} \quad (31)$$

We first transform I by introducing the tensor T from (14):

$$I = \int (\mathbf{r} \text{ div } \mathbf{T}) d\tau + 3 \int u \Delta u^* d\tau \quad (31a)$$

The first term reduces by integration by parts to

$$\begin{aligned} - \int \sum T_{ii} d\tau &= - \int \left\{ 2 \sum \frac{\partial u}{\partial x_i} \frac{\partial u^*}{\partial x_i} - 3 \sum \frac{\partial u}{\partial x_j} \frac{\partial u^*}{\partial x_j} \right\} d\tau \\ &= + \int \sum \frac{\partial u}{\partial x_i} \frac{\partial u^*}{\partial x_i} d\tau = - \int u \Delta u^* d\tau. \end{aligned}$$

If we introduce the second term from (31a) and take into account (8a) we get

$$I = 2 \int u \Delta u^* d\tau = - \frac{8\pi^2 m}{h^2} \cdot 2K \quad (31b)$$

We then transform the first integral in II by (29a):

$$\begin{aligned} \int \nabla u(\mathbf{r} \text{ grad } u^*) d\tau &= -3 \int \nabla u u^* d\tau - \int \nabla u^*(\mathbf{r} \text{ grad } u) \\ &\quad - \int u^* u(\mathbf{r} \text{ grad } \nabla) d\tau. \end{aligned}$$

The first two integrals on the right cancel out with the two last terms in the equation of definition (31) of II, and, by (26), there remains

$$\text{II} = \frac{8\pi^2 m}{\hbar^2} \int u^* u(\mathbf{r} \text{ grad } \nabla) d\tau = -\frac{8\pi^2 m}{\hbar^2} \text{R} \quad (31c)$$

Substitution of (31b, c) in (30) thus gives the analogue to the classical form (27) of the law of virial:

$$\dot{Q} = 2K + \text{R} \quad (32)$$

If, in particular, we are dealing with a stationary state in which u and u^* depend on the time in the manner given by

$$e^{\pm \frac{2\pi i}{\hbar} Et}$$

then Q becomes independent of t and $\dot{Q} = 0$. Thus whereas in classical mechanics we must explicitly average over the time in order to make the left-hand side of (27) vanish, the wave-mechanical method contains this averaging over the time implicitly and makes the left-hand side of (32) vanish spontaneously. Hence for stationary states we always have

$$\text{R} = -2K \quad (33)$$

If, further, all forces included in V are of electrical origin, then V becomes a homogeneous function of the -1^{th} degree in the co-ordinates, and Euler's theorem holds (cf. I, p. 548):

$$(\mathbf{r} \text{ grad } V) = -V.$$

From (26) and (2) it then follows that

$$\text{R} = \int u u^* \nabla d\tau = \text{U}.$$

Thus (33) states that the wave-mechanical mean of the kinetic energy is equal to half of the wave-mechanical mean of the potential energy taken with the reverse sign. The usefulness of the corresponding theorem in the older quantum theory is well known. In wave-mechanics we have also made use of it occasionally (cf. § 7, B, Chap. I). Note the parallelism between our present proof of the theorem and the former proof in I, Note 5, p. 547.

It is remarkable that the preceding calculations can be extended so that not *one* wave-state u is considered but *simultaneously* two states u_n and u_m . It is only necessary to replace $u u^*$ in forming the wave-

mechanical mean values by $u_n u_m^*$ or, respectively, u grad u^* by u_n grad u_m^* , etc. In the language of matrices this means that we can pass from the diagonal elements of the matrices in question directly to an arbitrary non-diagonal element.

E. Law of Areas in Differentials

Besides the *integral law of areas*, which we have treated under C there is also a very interesting *differential* or *specific law of areas*, which will be useful to us in the following section. We shall restrict ourselves to the simplest time-free wave-equation, which we shall write in the abbreviated form:

$$(\Delta + \lambda)\psi = 0, \quad \lambda = \frac{8\pi^2 m}{h^2}(E - V) \quad (34)$$

From this equation we form the moment with respect to an appropriately chosen point $r = 0$, that is, we apply to each member of (34) the "moment-operator"

$$M = [r \text{ grad}] \quad (35)$$

We get

$$M\Delta\psi + \lambda M\psi = -M\lambda \cdot \psi \quad (36)$$

Since $E = \text{constant}$,

$$-M\lambda = \frac{8\pi^2 m}{h^2}[r \text{ grad } V]$$

denotes the *non-integrated moment of the force* (except for a constant factor) in contrast with the previous *integral moment M of the force* of eqn. (17). We shall show that

$$M\Delta\psi = \Delta M\psi \quad (37)$$

holds, so that we may also write

$$(\Delta + \lambda)M\psi = -M\lambda \cdot \psi \quad (38)$$

This equation is particularly interesting in the case where V is a pure function of r and hence, cf. (24), $M\lambda$ becomes equal to zero. Then, by (38) $M\psi$ satisfies the same linear differential equation as ψ itself, namely the equation

$$(\Delta + \lambda)M\psi = 0 \quad (39)$$

Thus $M\psi$ must be built up in a certain sense linearly with constant coefficients from the integrals of (34), that is, from the proper functions ψ that belong to the same proper value E . This is the wave-mechanical equivalent of the *special law of areas*, that is, the conservation of the moment of momentum, being now in a differential and *not*, as above, in an integral form.

We have yet to supply the proof of (37). We first do this somewhat unsymmetrically in rectangular co-ordinates x, y, z ; it is sufficient to consider the component

$$M_{xy} = x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x}$$

We apply it separately to the three summands of $\Delta\psi$. In the case of $\partial^2\psi/\partial x^2$ we obviously get at once

$$M \frac{\partial^2 \psi}{\partial x^2} = \frac{\partial^2}{\partial x^2} M\psi.$$

On the other hand, we get in the case of the other two summands

$$M \frac{\partial^2 \psi}{\partial x^2} = x \frac{\partial^2 \psi}{\partial x^2 \partial y} - y \frac{\partial^2 \psi}{\partial x^2} = \frac{\partial^2}{\partial x^2} \left(x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} \right) - 2 \frac{\partial^2 \psi}{\partial x \partial y}. \quad (39a)$$

$$M \frac{\partial^2 \psi}{\partial y^2} = x \frac{\partial^2 \psi}{\partial y^2} - y \frac{\partial^2 \psi}{\partial x \partial y} = \frac{\partial^2}{\partial y^2} \left(x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} \right) + 2 \frac{\partial^2 \psi}{\partial x \partial y}. \quad (39b)$$

In the sum of these three expressions the additional term $\frac{\partial^2 \psi}{\partial x \partial y}$ actually drops out and we get, in agreement with (37): $M\Delta\psi = \Delta M\psi$.

Our proof is not quite satisfactory, as M is a vector, whereas the differential operator Δ is only to be applied to a scalar. This circumstance is known to be of no account in the case of rectangular co-ordinates, but is an obstacle where other co-ordinates, for example polar, are used. We therefore give in addition to (37) the general vectorial form, that is, the form valid for arbitrary co-ordinates:

$$M\Delta\psi = - \text{curl curl } M\psi. \quad (37a)$$

so that we get instead of (38)

$$(- \text{curl curl} + \lambda)M\psi = - M\lambda. \psi \quad (38a)$$

The proof of (37a) is most conveniently carried out by means of the symbolic ∇ -operator of ordinary vector-calculus. We have

$$\begin{aligned} \text{curl } M\psi &= [\nabla[r\nabla]]\psi = r(\nabla\nabla)\psi - \nabla(\nabla r)\psi \\ &= r\Delta\psi - 3 \text{ grad } \psi. \end{aligned}$$

If we once again take the "curl" here, the last term drops out and we get

$$\text{curl curl } M\psi = [\nabla r]\Delta\psi = - [r\nabla]\Delta\psi = - M\Delta\psi,$$

that is, (37a).

We apply eqn. (39) to the Kepler problem or, more generally, to a motion under the influence of arbitrary central forces that originate in a nucleus at the point $r = 0$. We conclude as above that $M\psi$ must be built up linearly with constant coefficients from the proper functions ψ that belong to the same E . (Only the *proper functions* come into question for this, that is, only the continuous solutions of the wave-equation, *not* the solutions that are discontinuous at the singular points, because $M\psi$ has the same properties of continuity as ψ .) In the case of the Kepler problem there belong to a proper value $E = E_n$ all proper functions that have the same principal quantum number n ; in the case of an atom with an arbitrary central force the degeneracy is less, the proper value $E = E_n$, depends on l as well as on n (separation of the Balmer terms into the

S-, P-, D- . . . terms). The degeneracy then affects only the third quantum number m , so that to a given E there belong only $2l + 1$ different proper functions $\psi = \psi_m$, which differ in the values of the suffix m . We must therefore assume $M\psi$ generally as a linear combination of these proper functions ψ_m :

$$M\psi = \sum C_m \psi_m \quad (40)$$

where the C_m 's are constants that must be determined for each individual case.

This is done most simply in the case of the xy -component of M ("the moment about the z -axis"). For here

$$M_{xy}\psi = x \frac{\partial \psi}{\partial y} - y \frac{\partial \psi}{\partial x} = \frac{\partial \psi}{\partial \phi} \quad (41)$$

holds, according to the remarks made in connection with eqn. (2) in § 10, Chap. II, so long as we choose the z -axis as the polar axis of an r, θ, ϕ -system. If, moreover, we write the proper function ψ , from which we start, in the form $\psi = R \cdot P_l^m(\cos \theta) e^{im\phi}$, then by (41) $M_{xy}\psi = im\psi$. So the sum in (40) in this case reduces to one term; we have $C_m = im$ and $C_{m'} = 0$ for all m' 's, which differ from the m of the initial function (*Ausgangsfunktion*).

On the other hand, we find for the other two components of M by means of calculations with spherical harmonics that we shall suppress here:

$$\left. \begin{aligned} M_{yz}\psi_m &= \left(-\sin \phi \frac{\partial}{\partial \theta} - \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right) \psi_m \\ &= -\frac{i}{2} \left[\psi_{m+1} + \{l(l+1) - (m-1)m\} \psi_{m-1} \right] \\ M_{zx}\psi_m &= \left(\cos \phi \frac{\partial}{\partial \theta} - \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right) \psi_m \\ &= -\frac{i}{2} \left[\psi_{m+1} - \{l(l+1) - (m-1)m\} \psi_{m-1} \right] \end{aligned} \right\} \quad (41a)$$

So the sum (40) in the present case consists of *two* terms.

It is clear that the process which brought us from (34) to (39) may be repeated. For if we again assume, for example, that $M\lambda = 0$, that is, $V = V(r)$, and take M_1, M_2 as standing for any two rectangular components of the moment-vector (35), we obviously have

$$M_1(\Delta + \lambda)M_2\psi = (\Delta + \lambda)M_1M_2\psi = 0.$$

Thus, in the sense of eqn. (40) $M_1M_2\psi$ must also allow itself to be built up linearly from the proper functions ψ of the same proper value. The same holds of a sum of such products M_1M_2 or of a product of more than two factors M .

We consider in particular, as a preparation for the following section,

the sum of the squares

$$M^2 = M_{xy}^2 + M_{yz}^2 + M_{zx}^2 \quad (42)$$

applied to any proper function of the Kepler problem with the azimuthal quantum number l , and assert that then the expression (40) resolves into the simple relation

$$M^2\psi = -l(l+1)\psi \quad (43)$$

For, by (41) and (41a) the meaning of the operator (42) is

$$\left. \begin{aligned} M^2 &= \frac{\partial^2}{\partial \phi^2} + \left(\sin \phi \frac{\partial}{\partial \theta} + \cos \phi \cot \theta \frac{\partial}{\partial \phi} \right)^2 \\ &\quad + \left(\cos \phi \frac{\partial}{\partial \theta} - \sin \phi \cot \theta \frac{\partial}{\partial \phi} \right)^2 \\ &= \frac{\partial^2}{\partial \theta^2} + \cot \theta \frac{\partial}{\partial \theta} + \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} \end{aligned} \right\} \quad (43a)$$

But by eqns. (1) *et seq.* of § 2, Chap. I, this is precisely the differential expression in the equation of the spherical harmonic functions (cf. p. 8); when applied to a spherical harmonic function of degree l , that is, to an aggregate of the form

$$S = \sum_{-l}^{+l} c_m P_l^m(\cos \theta) e^{im\phi}$$

it gives, by eqns. (1a) and (11) of § 2, Chap. I, the value $-l(l+1)S$. This proves eqn. (43).

If we further alter the meaning of M in (35) by adding the factor $\hbar/2\pi i$, which is consistent with the sense of the definition of the momentum-operator π in eqn. (8), that is, if we write in place of (35)

$$M = \frac{\hbar}{2\pi i} [\mathbf{r} \text{ grad}] \quad (44)$$

(43) becomes converted into

$$M^2\psi = l(l+1) \left(\frac{\hbar}{2\pi} \right)^2 \psi.$$

We can then say, if we extract the square root in a rather daring fashion, that

$$|M| = \sqrt{l(l+1)} \frac{\hbar}{2\pi} \quad (48b)$$

This formula replaces in a certain sense the representation of the orbital momentum in the old quantum theory

$$|M| = l \frac{\hbar}{2\pi}$$

and leads to the substitution of l^2 by $l(l+1)$ which we found necessary, for example in *Atombau*, Chap. VIII, p. 629, in order to be able to make the derivation of the anomalous Zeeman effect on the basis of the atomic model fit the experimentally confirmed g -formula of Landé.

F. The Uncertainty-Relation and Indefiniteness

Schrödinger has shown by using the oscillator as an example that in some circumstances a wave-group limited in space, that is, a "wave-packet," can hold together without appreciably diffusing. Some other simple forms of movements (free fall, circular motion and so forth) have also been investigated from the same point of view. The interest in these problems was probably aroused by the wish to build up the material point, for example, the electron, wave-mechanically and to reduce it to continuous distributions of field. We regard this objective, as already mentioned in the preface, as illusory. That is why in the preceding pages we have not enquired into the way in which the wave-pictures hold together in space but into their behaviour towards the fundamental laws of mechanics. No matter how much the wave-picture diffuses in the course of its motion, one point can be sharply defined that obeys the ordinary law of conservation of the centre of inertia. In the same way on the basis of the wave-picture a moment can be defined which under certain circumstances remains constant during the motion, and under other circumstances behaves just like the moment of momentum of the ordinary law of areas.

The questions of the conservation or flowing apart of wave-groups has been brought into a more general theoretical relationship by an important paper by Heisenberg.* To describe Heisenberg's ideas fully we should have to discuss the transformation theory of Jordan and Dirac,† which, with all other quantum-statistical questions, we wish to leave out of consideration. We therefore satisfy ourselves here with a few indications only.

We start out from the fundamental equation of quantum-mechanics in the form in which we set it up in eqn. (5), § 4, Chap. I,

$$pq - qp = \frac{h}{2\pi i} \quad . \quad . \quad . \quad . \quad . \quad (45)$$

We see that in (45) p , q cannot have the classical meaning of these quantities, for otherwise we should have zero on the right-hand side. So if we assume eqn. (45) to be correct—and we have no reason for doubting its validity—we must obviously revise our conceptions of locality and velocity.

* W. Heisenberg, Über den anschaulichen Inhalt der quanten-theoretischen Kinematik und Mechanik, *Zeitschr. f. Phys.*, **43**, 172 (1927); cf. also N. Bohr, *Naturwiss.*, **16**, 245 (1928).

† Cf. the references in Chap. I, p. 88.

To fix the position of an object, for example, an electron, we must perform experiments. Thus, following Heisenberg, we may choose to direct a "γ-ray microscope" on to the particle. The wave-length of the incident light is then very small and we can fix the position of the electron very accurately. For the width d of a structure just within the range of resolution for a wave-length λ of the light used is related to the numerical aperture ϵ of the instrument by the equation

$$d = \frac{\lambda}{2\epsilon} \quad . \quad . \quad . \quad . \quad . \quad (46)$$

If we now assume that the incident light is parallel, then we know, besides the quantity h/λ of the momentum of the incident light-quantum (photon), also its direction. On the other hand, the finite aperture of the instrument will prevent our being able to measure accurately the recoil that occurs owing to the Compton scattering of the light-quantum at the electron; we only get the momentum-component situated in the plane of the object after the collision with an error which is proportional to the size of the aperture and to the initial momentum of the light-quantum, that is, proportional to

$$2\epsilon\hbar/\lambda \quad . \quad . \quad . \quad . \quad . \quad . \quad (47)$$

Hence, by (47) and (46) the product of the accuracies p_1 and q_1 with which the momentum p and the position q of the electron can be determined becomes

$$p_1 q_1 = h \quad . \quad . \quad . \quad . \quad . \quad (48)$$

This is Heisenberg's "*relation of indefiniteness*." It is postulated as being *generally valid for any pair of canonically conjugate variables*. If we choose the energy and time as such a pair, then (48) teaches us directly that for stationary states in which the energy has a precisely fixed value, all statements about motions of the electron in time become meaningless. If one of the conjugate variables is fixed then in principle the other remains indefinite. This point was mentioned on pp. 84 and 192.

As Heisenberg shows, (48) is intimately connected with the transformation theory, due to Jordan and Dirac, of the probability amplitudes in the phase-space. To every particle there may be assigned a "probability-packet" (*Wahrscheinlichkeitspaket*), the region of which expresses the possibility of sojourn (*Aufenthaltsmöglichkeit*) of the particle; this "packet" spreads out in the course of time. If, say, we have fixed the position x of an electron with an accuracy Δx at the time t , then after the interval t_1 the probability-packet has spread out over a region in such a way that in a new measurement the electron will be encountered at every point of it with equal probability. The new measurement again contracts the range of play (*Spielraum*) to its original value Δx , and so forth. In the classical theory this was also the case in principle when the initial

state was insufficiently known. The only new features are the statistical character of the laws of motion and the restriction of the accuracy of a co-ordinate by its conjugate.*

But whereas in the classical theory the object was to render all laws objective and to release them as much as possible from the observing subject and the observing apparatus, this is no longer possible, according to Heisenberg and Dirac, now that theory and the means of observation have been refined; indeed, with Heisenberg we may regard the root of the quantum theory to lie in just that range of play in which according to theory the observation must become indefinite.

§ 10. Concerning the Nature of the Electron

It is difficult to report on the work of Dirac (cf. the reference in Chap. I, p. 118) in its present stage of development. But we cannot refrain from doing so, one reason being that his researches are of pre-eminent importance and another that they urgently call for a simplified treatment.

A. The Electron under No Forces

Dirac starts out from the relativistic wave-equation of the electron, which we have formulated in eqn. (7), § 9, Chap. I. He objects to this equation on the ground that it is untrue to its model in classical mechanics, the partial differential equation of Hamilton. With respect to the energy or the time differential quotient Hamilton's equation is of the *first* degree. Hence in the corresponding wave-equation we should also expect only a differential quotient of the *first* order with regard to t . But our relativistic wave-equation was of the *second* order both in t and in the space-co-ordinates. For it was

$$\sum_1^4 \{\Omega_k^2 - A^2\}u = 0 \quad . \quad . \quad . \quad (1)$$

Here, as there is no field in this case, Ω_k denotes, in consequence of eqn. (5a) of § 9, Chap. I, simply the differentiation with respect to x_k , and A , owing to eqn. (4) of § 9, Chap. I, signifies the abbreviation

$$A = \frac{\alpha E_0}{c^2} = \frac{2\pi}{hc} E_0 \quad . \quad . \quad . \quad (1a)$$

where α = the fine-structure constant, $E_0 = mc^2$ = the rest-energy,
 m = the rest-mass.

* According to E. H. Kennard, *Zeitschr. f. Phys.*, 44, p. 826 (1927), the relation (48) may be reduced to the more precise form

$$P_1 \Delta_1 > \frac{h}{2\pi}$$

The generalisation for a pair of arbitrary quantities is treated by E. V. Condon, *Science*, May, 1929, and Robertson, *Nature*, 1929.

On the basis of this postulate of Dirac we shall try to resolve eqn. (1) formally into two linear factors, of which one, when equated to zero, is then to represent the wave-equation of the (negative) electron.

It is true that it is not possible to resolve (1) into linear factors within the ordinary range of numbers even if we treat Ω_k , as is suggested by our way of writing (1), not as proper differential quotients but as differential operators. But we force a resolution if we introduce new numbers α_k whose laws of calculation have yet to be determined, and write in place of (1)

$$\left\{ \sum_1^4 \alpha_k \Omega_k - A \right\} \left\{ \sum_1^4 \alpha_k \Omega_k + A \right\} u = 0 \quad . \quad . \quad . \quad (2)$$

The α_k 's are to be constants, that is, quantities independent of the co-ordinate-values $x_1 \dots x_4$. Their laws of calculation result if we multiply out (2) and compare the powers of Ω with (1). We find in this way that

$$\alpha_k \alpha_l = 1 \dots l = k, \quad . \quad . \quad . \quad . \quad (3)$$

$$\alpha_k \alpha_l + \alpha_l \alpha_k = 0 \dots l \neq k \quad . \quad . \quad . \quad . \quad (4)$$

In setting up (4) we have taken into account that with the present meaning of Ω (in contrast with the following section) $\Omega_k \Omega_l = \Omega_l \Omega_k$; the left-hand side of (4) is the sum of the factors of both products $\Omega_k \Omega_l$ and $\Omega_l \Omega_k$. It must be equal to zero, because no such product occurs in (1). The form of (4) already shows us that our α 's are not commutative; for we have by (4) that

$$\alpha_k \alpha_l = - \alpha_l \alpha_k \dots l \neq k \quad . \quad . \quad . \quad . \quad (4a)$$

We now regard as the wave-equation of our electron under no forces the second factor in (2). [The first factor of (2) is only apparently different from it, for we can balance the change of sign of A with a common change of sign of the α 's, which is always possible by the equations of definition (3) and (4).] So we assert that the behaviour of the electron is described by the following *differential equation of the first order*

$$\left\{ \sum_1^4 \alpha_k \Omega_k + A \right\} u = 0 \quad . \quad . \quad . \quad . \quad (5)$$

whose coefficients α do not belong to the series of ordinary numbers.

In doing this we have really overshot our objective. We desired to have a differential equation which was to be of the first degree only in t , that is, in the relativistic co-ordinate x_4 . Instead of this we have found a differential equation which is of the first degree also in x_1, x_2, x_3 . This contradicts the model of the classical Hamiltonian equation in the other direction. For the Hamiltonian equation is of the second degree in the differential quotients with respect to the space-co-ordinates, and hence would lead us to expect a wave-equation of the second order in x_1, x_2, x_3 .

Hence we should probably regard the way that led us to (5) only as a heuristic method.

We have differed from Dirac in that we have added * the co-efficients α_k in a *symmetrical manner* to the four operators Ω_k , whereas Dirac distinguishes Ω_4 , the differentiation with respect to the time (he writes p_0 for Ω_4) and gives it the coefficient 1; on the other hand Dirac attaches his four coefficients (which he calls $\alpha_1, \alpha_2, \alpha_3$ and β) to the quantities $\Omega_1, \Omega_2, \Omega_3$ (or p_1, p_2, p_3 respectively) and Λ . This can make no material difference. We shall see that our symmetrical choice of the α 's offers certain advantages.

We shall postpone the detailed investigation of the α 's till section G. It will be sufficient to obtain a special solution, and one which is as convenient as possible, of eqns. (3) and (4) without our requiring to discuss the question of the general form of the solution.

B. The Electron in the Electromagnetic Field and its Magnetic Moment

We describe the field in terms of the 4-vector Φ of p. 100,

$$\Phi_1, \Phi_2, \Phi_3 = \frac{2\pi e}{hc} \mathbf{A}, \quad \Phi_4 = \frac{2\pi i e}{hc} \phi = \frac{2\pi i}{hc} V \quad (6)$$

where \mathbf{A} and ϕ are the ordinary vector- and scalar-potential respectively, and we now define Ω (cf. p. 101) by

$$\Omega_k = \frac{\partial}{\partial x_k} + i\Phi_k \quad (6a)$$

The equation of the electron is to retain exactly the form (5), the significance of the coefficients α being unchanged so that, as before, they are to be constant quantities and, in particular, independent of the field. We enquire what equation arises from the multiplication projected in (2), if the Ω 's now have the meaning (6), that is, depend on the field.

Instead of (2) we now write more simply

$$\left\{ \sum_1^4 \alpha_k \Omega_k \cdot \sum_1^4 \alpha_k \Omega_k - \Lambda^2 \right\} u = 0 \quad (2a)$$

To multiply out the two sums let us firstly consider two terms with the same summation-index :

$$\alpha_k \Omega_k \cdot \alpha_k \Omega_k \quad (7)$$

and, secondly, let us combine members of unequal summation indices k and l :

$$\alpha_k \Omega_k \cdot \alpha_l \Omega_l + \alpha_l \Omega_l \cdot \alpha_k \Omega_k \quad (8)$$

* The same is done by F. Möglich, *Zeitschr. f. Phys.*, 48, 852; J. V. Neumann, *ibid.*, 48, 866; H. Tetrode, *ibid.*, 48, 868 (1928).

The α 's being constant quantities may be commuted with the Ω 's; instead of (7) we may therefore also write

$$\alpha_i^2 \Omega_k \Omega_k = \Omega_k^2 \quad [\text{on account of (3)}] \quad (7a)$$

For the same reason we may write in place of (8)

$$\alpha_k \alpha_i \Omega_k \Omega_i + \alpha_i \alpha_k \Omega_i \Omega_k = \alpha_k \alpha_i (\Omega_k \Omega_i - \Omega_i \Omega_k) \quad [\text{on account of (4a)}].$$

Now, by (6a),

$$\begin{aligned} \Omega_k \Omega_i u &= \left(\frac{\partial}{\partial x_k} + i\Phi_k \right) \left(\frac{\partial}{\partial x_i} + i\Phi_i \right) u \\ &= \frac{\partial^2 u}{\partial x_k \partial x_i} + i \frac{\partial(\Phi_i u)}{\partial x_k} + i\Phi_k \frac{\partial u}{\partial x_i} - \Phi_k \Phi_i u. \end{aligned}$$

If we form $\Omega_i \Omega_k u$ in the same way, then in the expression for the difference the first and the last terms cancel; but the middle terms only partly cancel. For there remains

$$(\Omega_k \Omega_i - \Omega_i \Omega_k) u = i \left(\frac{\partial \Phi_i}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_i} \right) u \quad (8b)$$

The bracket on the right is the "four-dimensional curl" of the four-vector Φ . According to the well-known rules of the theory of electrons this is connected, on account of (6), with the field-quantities \mathbf{E} and \mathbf{H} in the following way:

for $k, l = 1, 2, 3$:

$$\frac{\partial \Phi_l}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_l} = \frac{2\pi e}{hc} \text{curl}_{kl} \mathbf{A} = \frac{2\pi e}{hc} \mathbf{H}_{kl} \quad (9a)$$

for $l = 4, k = 1, 2, 3$:

$$\frac{\partial \Phi_4}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_4} = \frac{2\pi ie}{hc} \left(\text{grad}_k \phi + \frac{1}{c} \mathbf{A}_k \right) = - \frac{2\pi ie}{hc} \mathbf{E}_k \quad (9b)$$

Our method of writing the quantities \mathbf{H} and curl with two indices is fully justified since, in contrast with \mathbf{E} and grad, they are not polar, but axial quantities, not distance but surface quantities.

If we collect together all terms (7) and (8) we get from (2a) by (7a) and (8a, b)

$$\left(\sum_1^4 \Omega_k^2 - \mathbf{A}^2 \right) u + \sum_{k,l} i \alpha_k \alpha_l \left(\frac{\partial \Phi_l}{\partial x_k} - \frac{\partial \Phi_k}{\partial x_l} \right) u = 0 \quad (10)$$

The second sum comprises not $4 \cdot 3 = 12$ terms but only six, since in our method of derivation every number-pair k, l is to be taken only once. We arrange these six terms in groups of three, corresponding to (9a) and

(9b), namely

$$\frac{2\pi ie}{\hbar c}(\alpha_1\alpha_2\mathbf{H}_{12} + \alpha_2\alpha_3\mathbf{H}_{23} + \alpha_3\alpha_1\mathbf{H}_{31})u \quad (10a)$$

$$\frac{2\pi e}{\hbar c}\alpha_4(\alpha_1\mathbf{E}_1 + \alpha_2\mathbf{E}_2 + \alpha_3\mathbf{E}_3)u \quad (10b)$$

Hence, according to Dirac, our earlier relativistic wave-equation (1) is to be supplemented by these expressions as correction terms. As they are attached to u they become added in a certain sense to the potential energy V . We recall, for example, the two simplest (non-relativistic) forms of the wave-equation:

$$\Delta\psi + \frac{8\pi^2m}{\hbar^2}(E - V)\psi = 0 \quad [\text{eqn. (11), § 1, Chap. I},]$$

$$\text{and } \Delta u - \frac{4\pi im}{\hbar} \frac{\partial u}{\partial t} - \frac{8\pi^2m}{\hbar^2}Vu = 0 \quad [\text{eqn. (13), § 5, Chap. I}].$$

A comparison, particularly of the last equation with (10a, b), shows that we must divide the expressions (10a, b) by $-8\pi^2mu/\hbar^2$, in order to make them comparable with V . In this way we obtain quantities that we shall call V_m and V_e as contributions to the potential energy:

$$V_m = \frac{e}{m} \frac{\hbar}{4\pi c} (-i)(\alpha_1\alpha_2\mathbf{H}_{12} + \alpha_2\alpha_3\mathbf{H}_{23} + \alpha_3\alpha_1\mathbf{H}_{31}) \quad (11a)$$

$$V_e = \frac{e}{m} \frac{\hbar}{4\pi c} (-\alpha_4)(\alpha_1\mathbf{E}_1 + \alpha_2\mathbf{E}_2 + \alpha_3\mathbf{E}_3) \quad (11b)$$

Our immediate interest is with the first factor

$$\mu = \frac{e}{m} \frac{\hbar}{4\pi c} \quad (12)$$

It denotes the value of the Bohr magneton, I, p. 249. Although we have assumed nothing about a magnetic axis or a magnetic moment of an electron we thus find that the characteristic expression of a magnetic moment, derived in a purely formal way, here presents itself, and to an amount exactly equal to that of a whole magneton such as the hypothesis of Goudsmit and Uhlenbeck demands for the electron. We also have the same factor in V_e , the electric energy, as in V_m , the magnetic energy; this means that the electron also has an electric moment of the same amount, which corresponds with an earlier postulate by J. Frenkel,* who, for reasons connected with relativistic invariance, ascribes to the electron a moment of the character of a six-vector, with three real magnetic and three imaginary electric components.

If this were a question of a process of adjustment of an ordinary kind, in which the magnetic axis of the electron formed the direction

* *Zeitschr. f. Phys.*, **37**, 243 (1926); **47**, 736 (1928).

cosines $\alpha_x, \alpha_y, \alpha_z$ with the co-ordinate axes, the magnetic energy of the electron in the magnetic field \mathbf{H} would be given by

$$V_m = (\vec{\mu}, \mathbf{H}) = \mu(\alpha_x \mathbf{H}_x + \alpha_y \mathbf{H}_y + \alpha_z \mathbf{H}_z).$$

According to this α_z corresponds to our factor $(-i)\alpha_1\alpha_2$ in eqn. (11a); corresponding results hold for α_x, α_y and for the position into which the electric moment adjusts itself.

We must, however, emphasise that the interpretation of the supplementary terms (10a, b) as magnetic and electric moments has a historical rather than a real foundation, as is shown by its derivation from the differential equation of the second order (2a). The true differential equation of the problem, which according to Dirac is of the first order, does not lead directly to these moments. In agreement with this we shall actually in the sequel not have to deal with them explicitly.

C. The Mechanical Moment of the Electron

We next show that the "law of areas" does not hold in the form in which we developed it in the section 9, E. For this purpose we shall assume that, as in the case of the hydrogen atom, the electron is situated in the field of a central force, so that the potentials (6) assume the values

$$\Phi_1 = \Phi_2 = \Phi_3 = 0, \quad \Phi_4 = \frac{2\pi i}{hc} V(r) \quad . \quad . \quad . \quad (13)$$

We take M to stand for any spatial component of the momentum-vector, for example,

$$M_{12} = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1},$$

for which we wrote M_{xy} in § 9. Applied to Ω_1 or Ω_2 these components give, similarly to the results of eqns. (39a, b) of § 9,

$$\begin{aligned} M_{12}\Omega_1 &= \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\right) \frac{\partial}{\partial x_1} = \frac{\partial}{\partial x_1} \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\right) - \frac{\partial}{\partial x_2} \\ &= \Omega_1 M_{12} - \Omega_2, \\ M_{12}\Omega_2 &= \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\right) \frac{\partial}{\partial x_2} = \frac{\partial}{\partial x_2} \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}\right) + \frac{\partial}{\partial x_1} \\ &= \Omega_2 M_{12} + \Omega_1. \end{aligned}$$

Further, we clearly have $M_{12}\Omega_3 = \Omega_3 M_{12}$ and, also, on account of the vanishing of $M_{13}V(r)$, $M_{13}\Omega_4 = \Omega_4 M_{13}$. Thus if we denote the linear differential expression on the left-hand side of eqn. (5) by the abbreviation L , we have

$$M_{12}L(u) = LM_{12}(u) + (\alpha_2\Omega_1 - \alpha_1\Omega_2)u \quad . \quad . \quad . \quad (14)$$

In contrast with eqns. (39a, b) of § 9 the supplementary terms arising from the differentiation do not now cancel. One reason for this is that our differential equation is now of the first order and not, as before, of the

second order; another reason is that now our coefficients α_1, α_2 have inserted themselves. We may formulate this result by saying:

The law of areas in its ordinary form, representing the conservation of the moment of momentum is not obeyed in the case of the electron in spite of our assumption of a central field of force.

But we can restore the law of areas if we change the operator M of the momentum to

$$N = M + \beta \quad (15)$$

where β is to be a number of the same character as our coefficients α , to which we shall presently assign convenient values. Just as the α 's are not commutative among themselves (cannot be "interchanged") so β is not interchangeable with the α 's; accordingly $\beta L \neq L\beta$. We next form

$$NL - LN = ML - LM + \beta L - L\beta \quad (15a)$$

and consider, in particular, the component N_{12} . We then have, if we take (14) into account,

$$N_{12}L - LN_{12} = \alpha_2\Omega_1 - \alpha_1\Omega_2 + \beta L - L\beta \quad (16)$$

If we now choose β so that

$$\alpha_2\Omega_1 - \alpha_1\Omega_2 + \beta L - L\beta = 0 \quad (17)$$

then eqn. (16) asserts that

$$N_{12}L = LN_{12} \quad (18)$$

And the same equation holds for every other purely spatial component of N , that is, for any two indices k and l of the series 1, 2, 3:

$$N_{kl}L = LN_{kl} \quad k, l = 1, 2, 3 \quad (18a)$$

so long as we only determine β from the equation of condition

$$\alpha_l\Omega_k - \alpha_k\Omega_l + \beta L - L\beta = 0 \quad (17a)$$

which is analogous to (17).

The fact that the eqn. (18a) does not also hold for the space-time components, for example, for $k = 1, l = 4$, is due to our having used the relation $M_k V(r) = [r \text{ grad}] V(r) = 0$ in deriving it, which is true only so long as r denotes the spatial radius-vector.

We shall see that the eqns. (18) and (18a) respectively already contain the fact of electron spin. First we shall deal with the resolution of eqn. (17) in terms of β . We write (17) explicitly as

$$\alpha_2\Omega_1 - \alpha_1\Omega_2 + \beta \sum_1^4 \alpha_i\Omega_i - \sum_1^4 \alpha_i\Omega_i\beta \quad (19)$$

In this expression we have twice omitted the term in L involving A since A is an ordinary number and hence, certainly, $\beta A - A\beta = 0$. In (19) the coefficients of all four Ω 's must individually equal zero. So we obtain four equations for determining β :

$$\begin{cases} \alpha_2 + \beta\alpha_1 - \alpha_1\beta = 0 \\ -\alpha_1 + \beta\alpha_2 - \alpha_2\beta = 0 \end{cases} \quad (20a) \quad \begin{cases} \beta\alpha_3 - \alpha_3\beta = 0 \\ \beta\alpha_4 - \alpha_4\beta = 0 \end{cases} \quad (20b)$$

According to (20b) β is commutative with α_3 and α_4 . This is certainly the case if we set

$$\beta = c\alpha_1\alpha_2 \quad (21)$$

where c is an ordinary number.

By applying the commutative rule (4a) we actually get

$$\beta\alpha_3 = c\alpha_1\alpha_2\alpha_3 = -c\alpha_1\alpha_3\alpha_2 = +c\alpha_3\alpha_1\alpha_2 = \alpha_3c\alpha_1\alpha_2 = \alpha_3\beta.$$

The first eqn. of (20b) is therefore fulfilled by our assumption (21). The same holds for the second eqn. of (20b).

If on the other hand we insert eqn. (21) in the first eqn. (20a) we get

$$\alpha_2 + c\alpha_1\alpha_2\alpha_1 - c\alpha_1\alpha_1\alpha_2 = 0,$$

or, in view of (3) and (4a),

$$\alpha_2 - c\alpha_2\alpha_1^2 - c\alpha_1^2\alpha_2 = \alpha_2 - 2c\alpha_2 = 0,$$

that is,

$$c = \frac{1}{2} \quad (21a)$$

But the second eqn. of (20a) leads to the same value of c ; for by similar calculations we get

$$\begin{aligned} -\alpha_1 + c\alpha_1\alpha_2\alpha_2 - c\alpha_2\alpha_1\alpha_2 &= -\alpha_1 + c\alpha_1\alpha_2^2 + c\alpha_2^2\alpha_1 \\ &= -\alpha_1 + 2c\alpha_1. \end{aligned}$$

We thus have, if we at once generalise (21) for arbitrary indices, $k, l = 1, 2, 3$, as a possible solution of eqn. (17a),

$$\beta = \frac{1}{2}\alpha_k\alpha_l \quad (22)$$

and for the symbol (15),

$$N_{kl} = M_{kl} + \frac{1}{2}\alpha_k\alpha_l.$$

As at the end of § 9, E, we here wish to make a correction in the definitions of N and M by taking up in these definitions the appropriate factor $\hbar/2\pi i$. From the preceding equation there then arises

$$N_{kl} = M_{kl} - \frac{1}{2} \frac{\hbar}{2\pi} i \alpha_k \alpha_l \quad (23)$$

which we supplement with the corresponding equation of definition of M_{kl} ,

$$M_{kl} = \frac{\hbar}{2\pi i} \left(x_k \frac{\partial}{\partial x_l} - x_l \frac{\partial}{\partial x_k} \right) \quad (23a)$$

We now revert to eqn. (18a) and apply it to a solution of the eqn. $Lu = 0$. On account of the vanishing of the left-hand side (18a) then states that, besides u , $N_{kl}u$ is also an integral of $L = 0$. From this we infer as in § 9, E, that

$$N_{kl}u = Cu \quad (24)$$

or more generally an aggregate $\Sigma C_{kl}u_l$ of solutions u_l , which satisfy the same differential equation $L = 0$.

Introducing polar co-ordinates r, θ, ϕ we assume for u (with the usual special form of dependence on the time) the form

$$u = R\Theta\Phi e^{\frac{2\pi i}{h}Et} \quad (25)$$

and choose $k = 1, l = 2$.

We then get [cf. eqn. (41), § 9, and our above convention about the factor $\hbar/2\pi i$ that is to be added]:

$$M_{12}u = \frac{\hbar}{2\pi i} \frac{\partial u}{\partial \phi},$$

and our equations (23) and (24) state [it is sufficient in this case to use eqn. (24), that is, the single-termed form of the right-hand side]:

$$\frac{\hbar}{2\pi i} \frac{\partial u}{\partial \phi} - \frac{1}{2} \frac{\hbar}{2\pi} i \alpha_1 \alpha_2 u = Cu \quad (26)$$

Thus u satisfies a linear differential equation in ϕ with constant co-efficients; such an equation is generally integrated by means of the exponential function. Hence we are justified in setting

$$\Phi = e^{im\phi} \quad (26a)$$

where, on account of the one-valued character of u , m must be an integer. Consequently eqn. (26) runs

$$\frac{\hbar}{2\pi} \left(m - \frac{1}{2} i \alpha_1 \alpha_2 \right) = C \quad (27)$$

It determines the value of the constant of integration C . Further it tells us that C has the significance of an areal constant, that is, of the moment of momentum for the z -axis (the axis $\theta = 0$). This moment of momentum has not, as in the earlier theory, the value $mh/2\pi$ but is supplemented by a moment of momentum of the amount

$$\frac{1}{2} \frac{\hbar}{2\pi}$$

which we must interpret as *electron spin*. The factor $-i\alpha_1\alpha_2$ in (27) is, as we have already remarked in connexion with eqn. (11a), a sort of direction cosine, namely the cosine of the inclination of the magnetic axis of the electron to the z -axis. We may therefore say: *the inclination of the electron spin coincides with the inclination of the magnetic axis*. But whereas the value of the magnetic moment is that of a complete magneton, the mechanical moment is equal to only half the quantum unit $\hbar/2\pi$ of the moment of momentum.

In this way the hypothesis of Goudsmit and Uhlenbeck is derived as regards not only its magnetic but also its mechanical moment in a *purely formal way*, without our having had to base our argument on an arbitrary model concerning the structure or the motion of the electron. So this

hypothesis loses its peculiar character and is placed on a firm mathematical foundation.

What we have here said about the z -axis may, of course, be extended to every other axis fixed in space, so long as, in given cases, the generalisation ($\Sigma C_k u_k$ instead of Cu) provided for in (24) is taken into account. The fact of our favouring the z -axis is due to the particular choice of the polar co-ordinate system, which distinguishes the z -axis as the polar axis.

D. Transformation to Polar Co-ordinates

Our original differential equation is, in rectangular co-ordinates if we insert the meaning of Ω from (6a) and (13), and that of A from (1a), and eliminate the dependence on time by (25):

$$\left\{ \sum_1^3 \alpha_k \frac{\partial}{\partial x_k} + \frac{2\pi}{hc} [\alpha_4(E - V) + E_0] \right\} u = 0 \quad (28)$$

To transform it to polar co-ordinates r, θ, ϕ we follow Dirac by introducing the quantity

$$\gamma = \sum_1^3 \alpha_k \frac{x_k}{r} \quad (29)$$

γ is a function of the ratios of the co-ordinates $x_1 : x_2 : x_3$, that is, independent of r and dependent only on θ and ϕ , but, on account of the occurrence of the α 's, not a function of the ordinary continuum of numbers. Obviously $\gamma^2 = 1$. For from (29) it follows that

$$\gamma^2 = \sum_1^3 \alpha_k \frac{x_k}{r} \sum_1^3 \alpha_l \frac{x_l}{r},$$

that is, if we write the terms $k = l$ and $k \neq l$ separately,

$$\gamma^2 = \sum_{k=1}^3 \alpha_k^2 \frac{x_k^2}{r^2} + \sum_{k,l=1}^3 (\alpha_k \alpha_l + \alpha_l \alpha_k) \frac{x_k x_l}{r^2},$$

and on account of (3) and (4) this actually gives

$$\gamma^2 = 1 \quad (30)$$

We next transform the first term of (28) by multiplying it by $\gamma^2 = 1$. This gives

$$\gamma \sum_1^3 \alpha_k \frac{x_k}{r} \sum_1^3 \alpha_l \frac{\partial u}{\partial x_l} \quad (31)$$

We again separate the terms with $l = k$ from those with $l \neq k$. On account of $\alpha_k^2 = 1$ the former give

$$\frac{\gamma}{r} \sum_1^3 x_k \frac{\partial u}{\partial x_k} = \gamma \frac{\partial u}{\partial r} \quad (32)$$

In view of (23a) the latter give

$$\frac{\gamma}{r} \sum \alpha_k \alpha_l \left(x_k \frac{\partial u}{\partial x_l} - x_l \frac{\partial u}{\partial x_k} \right) = \frac{\gamma}{r} \frac{2\pi i}{h} \sum \alpha_k \alpha_l M_{kl} u. \quad (33)$$

in which the pair of indices k, l run through the values 1, 2, or 2, 3 or 3, 1. We here introduce the important abbreviation

$$X = \frac{2\pi i}{h} (\alpha_1 \alpha_2 M_{12} + \alpha_2 \alpha_3 M_{23} + \alpha_3 \alpha_1 M_{31}) \quad (34)$$

In virtue of the quantities M , X is a differential operator, but in view of the factors α it is not an operator that can be applied in the range of ordinary numbers.

Hence our expression (31) is as a whole

$$\gamma \left(\frac{\partial u}{\partial r} + \frac{Xu}{r} \right) \quad (35)$$

Since this represents only an identical transformation of the first term of eqn. (28), (28) is identical with

$$\left\{ \gamma \left(\frac{\partial}{\partial r} + \frac{X}{r} \right) + \frac{2\pi}{hc} [\alpha_4 (E - V) + E_0] \right\} u = 0 \quad (36)$$

This equation depends explicitly on the variable r only; the variables θ and ϕ are contained implicitly in the differential operator X and in the quantity γ defined by (29).

E. Introduction of Matrices with Two Rows

We now pass on to the formal study of our coefficients α . We proceed in steps, beginning with the trivial case of *one* quantity α and *one* variable x . The equations of condition (3, 4) then reduce to the one equation $\alpha^2 = 1$, $\alpha = \pm 1$. If we regard α as an operator that is to be performed on x , $\alpha x = +x$ denotes *identity*, $\alpha x = -x$ denotes *reflection at the origin*. The first operation is of no interest, the second produces from the point P the symmetrical point P_1 (cf. Fig. 30a).

We next consider the case of *two* operators, which we shall call ϵ_1, ϵ_2 to distinguish them from the above α , and *two* variables x_1, x_2 . The condition (4) runs

$$\epsilon_1 \epsilon_2 = -\epsilon_2 \epsilon_1 \quad (37)$$

So we look for two operations that are "anti-commutative" that is, such that when their order is changed the result is reversed. The

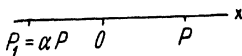


FIG. 30a.

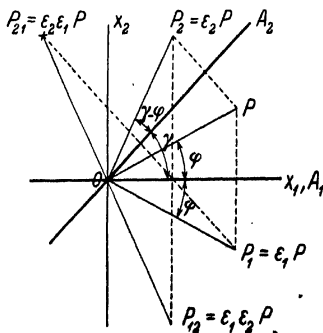


FIG. 30b.

conditions (3), that is, $\epsilon_1^2 = \epsilon_2^2 = 1$, are of no account here since they allow themselves to be fulfilled afterwards by normalising the two operations (addition of a scalar factor to each), without the condition (37) being affected.

We choose ϵ_1 and ϵ_2 as reflections, ϵ_1 as the reflection at an axis A_1 for which we may take the x_1 -axis, ϵ_2 as the reflection at an axis A_2 , which we shall suppose forms the angle γ with A_1 (cf. Fig. 30b). ϵ_1 transforms any arbitrarily taken point P into P_1 , ϵ_2 transforms the same point into P_2 . Let the point P_{12} correspond similarly to the combined operation $\epsilon_1\epsilon_2$ (first ϵ_2 and then ϵ_1), and the point P_{21} to the combination $\epsilon_2\epsilon_1$. We write down the angles at which these points lie, counting from the x_1 -axis and assuming that ϕ is the angle corresponding to the initial point P :

$$\begin{array}{ccccccccc} P & P_1 & & P_2 & & P_{12} & & P_{21} \\ \phi & -\phi & \phi & \phi + 2(\gamma - \phi) & -\phi & -2(\gamma - \phi) & -\phi & \phi + 2(\gamma + \phi) \end{array}$$

The condition (37) is equivalent to the condition that the points P_{12} and P_{21} should lie diametrically with respect to each other, that is, their angles are to differ by π . This means that

$$-\phi + 2(\gamma + \phi) + \phi + 2(\gamma - \phi) = 4\gamma = \pm \pi.$$

Thus the axes A_1 and A_2 must form an angle of 45° with each other. Our result is unique except for a common rotation of the axes in their plane and irrespective of which of the two we call A_1 or A_2 . The conditions $\epsilon_1^2 = \epsilon_2^2 = 1$ are clearly also fulfilled in our reflections, so that in this case it is unnecessary to add particular normalising factors.

Our result recalls the characters of symmetry in crystallography. When we speak of a "two-dimensional cubic" crystal, that is, of a square, its elements of symmetry are reflection at the sides and reflection at the diagonals.

We next write down the *linear transformations* that correspond to the operations ϵ_1 and ϵ_2 . Let x_1, x_2 be the co-ordinates of the initial point P , and x_1' and x_2' those of the transformed point P_1 or P_2 respectively. We clearly have the following, expressed in the usual diagrammatic way,

$$\epsilon_1 \left\{ \begin{array}{c|cc} & x_1 & x_2 \\ \hline x_1' & 1 & 0 \\ x_2' & 0 & -1 \end{array} \right. \quad \epsilon_2 \left\{ \begin{array}{c|cc} & x_1 & x_2 \\ \hline x_1' & 0 & 1 \\ x_2' & 1 & 0 \end{array} \right.$$

Which scheme corresponds to the order $\epsilon_1\epsilon_2$ and which to $\epsilon_2\epsilon_1$? For the sake of generality we first investigate this for the general transformations

$$A \left\{ \begin{array}{c|cc} & x_1 & x_2 \\ \hline x_1' & a_{11} & a_{12} \\ x_2' & a_{21} & a_{22} \end{array} \right. \quad B \left\{ \begin{array}{c|cc} & x_1 & x_2 \\ \hline x_1' & b_{11} & b_{12} \\ x_2' & b_{21} & b_{22} \end{array} \right.$$

To form AB we therefore apply the operation A to the point $x_1'x_2'$ which has already been transformed by B , and we obtain a point $x_1''x_2''$, namely:

$$\begin{aligned}
 x_1'' &= a_{11}x_1' + a_{12}x_2' = a_{11}(b_{11}x_1 + b_{12}x_2) + a_{12}(b_{21}x_1 + b_{22}x_2) \\
 &= (a_{11}b_{11} + a_{12}b_{21})x_1 + (a_{11}b_{12} + a_{12}b_{22})x_2 \\
 x_2'' &= a_{21}x_1' + a_{22}x_2' = a_{21}(b_{11}x_1 + b_{12}x_2) + a_{22}(b_{21}x_1 + b_{22}x_2) \\
 &= (a_{21}b_{11} + a_{22}b_{21})x_1 + (a_{21}b_{12} + a_{22}b_{22})x_2
 \end{aligned}$$

The result is the "multiplication of matrices" defined on p. 31, but here taken over from schemes of an infinite number of terms and applied to quadratic schemes. Thus the "composite matrix" AB is formed according to the rule: a row of A multiplied by a column of B and is represented by the scheme ($k = 1, 2$):

$$AB \left\{ \begin{array}{c|cc} & x_1 & x_2 \\ \hline x_1'' & \Sigma a_{1k}b_{k1} & \Sigma a_{1k}b_{k2} \\ x_2'' & \Sigma a_{2k}b_{k1} & \Sigma a_{2k}b_{k2} \end{array} \right.$$

Applied to our operations ϵ_1 and ϵ_2 the rule obviously gives, if we leave out the variables x from now on,

$$\epsilon_1\epsilon_2 = \begin{vmatrix} 0 & 1 \\ -1 & 0 \end{vmatrix}, \quad \epsilon_2\epsilon_1 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = -\epsilon_1\epsilon_2 \quad (38)$$

as eqn. (37) demands and as we have already established in connexion with Fig. 30b.

In addition to the two fundamental operations ϵ_1 , ϵ_2 we take the operation $\epsilon_1\epsilon_2$ derived from them, after having normalised it precisely like ϵ_1 and ϵ_2 . Thus we set

$$\epsilon_3 = c\epsilon_1\epsilon_2$$

and determine c from the condition $\epsilon_3^2 = 1$. This clearly gives $c^2 = -1$, $c = i$; for on account of (37) we have

$$\epsilon_1\epsilon_2\epsilon_1\epsilon_2 = -\epsilon_1^2\epsilon_2^2 = -1.$$

Thus our three operations, with this normalisation of ϵ_3 and in view of (38), run,

$$\epsilon_1 = \begin{vmatrix} 1 & 0 \\ 0 & -1 \end{vmatrix}, \quad \epsilon_2 = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix}, \quad \epsilon_3 = \begin{vmatrix} 0 & i \\ -i & 0 \end{vmatrix} \quad (39)$$

They not only fulfil the condition $\epsilon_k^2 = 1$, but also for *every* pair of indices k, l they also fulfil the condition

$$\epsilon_k\epsilon_l = -\epsilon_l\epsilon_k \quad (39a)$$

This is shown either by working out the corresponding matrices or more simply on the basis of the definition of ϵ_3 by the following equations:

$$\begin{cases} \epsilon_1\epsilon_3 = i\epsilon_1\epsilon_1\epsilon_2 = +i\epsilon_2, \\ \epsilon_3\epsilon_1 = i\epsilon_1\epsilon_2\epsilon_1 = -i\epsilon_2\epsilon_1\epsilon_1 = -i\epsilon_2, \\ \epsilon_2\epsilon_3 = i\epsilon_2\epsilon_1\epsilon_2 = -i\epsilon_1\epsilon_2\epsilon_2 = -i\epsilon_1, \\ \epsilon_3\epsilon_2 = i\epsilon_1\epsilon_2\epsilon_2 = +i\epsilon_1. \end{cases}$$

It is interesting also to consider the quantities normalised to -1 , although they differ from those normalised to $+1$ only unessentially,

namely by the factor i . We call them $\delta_1, \delta_2, \delta_3$ and have

$$\delta_1 = i\epsilon_1 = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix}, \delta_2 = i\epsilon_2 = \begin{vmatrix} 0 & i \\ i & 0 \end{vmatrix}, \delta_3 = i\epsilon_3 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}. \quad (40)$$

Like the ϵ 's they also behave anti-commutatively. But the following cyclic relationships also hold between them:

$$\delta_1\delta_2 = \delta_3, \quad \delta_2\delta_3 = \delta_1, \quad \delta_3\delta_1 = \delta_2. \quad (41)$$

We recognise this from the relationships between the δ 's and ϵ 's:

$$\begin{aligned} \delta_1\delta_2 &= i^2\epsilon_1\epsilon_2 = i\epsilon_3 = \delta_3, \\ \delta_2\delta_3 &= i^2\epsilon_2\epsilon_3 = i\epsilon_1 = \delta_1, \\ \delta_3\delta_1 &= i^2\epsilon_3\epsilon_1 = i\epsilon_2 = \delta_2. \end{aligned}$$

Consequently the δ 's become identical with the units i, j, k of the theory of quaternions introduced by Hamilton (the double use of i as a quaternion unit and as the imaginary root of -1 is unavoidable here). Actually, the quaternion units are defined by Hamilton by the relations

$$ij = k, \quad jk = i, \quad ki = j,$$

which are equivalent to (41); to these there are added the normalising conditions $i^2 = j^2 = k^2 = -1$ and the commutation rules $ij = -ji$, etc., in complete agreement with our rules of calculation for the δ 's.

The quaternion units and likewise our quantities δ form, when supplemented by the real unit ± 1 , a closed group of four elements which reproduces itself after arbitrary multiplications among its elements; it forms a "four-group." In the case of our ϵ 's this holds only when they are extended not only by the real units but also by the imaginary units $\pm i$, on account of their different normalisation.

The quaternion units have already been introduced into the theory of electrons by Pauli. Linking up with this, Dirac elaborated Pauli's two-dimensional matrices to four-dimensional matrices. We shall return to this point later. For the present our two-dimensional matrices, which are due to Pauli, are sufficient.

Our matrices ϵ are Hermitean in character [that is, are conjugate imaginaries with respect to the principal diagonal, cf. ϵ_3 in (39)]. It is interesting to note that in the region of three variables x_1, x_2, x_3 matrices of this character, which are at the same time anti-commutative, cannot exist.* The existence of four-dimensional matrices and the fact that they can be reduced to our two-dimensional matrices is clearly due to the circumstance that $4 = 2 \cdot 2$.

F. Integration in Angular Co-ordinates

We can put our law of areas to still further use than we have already done in section C. As mentioned in § 9, E, the moment operator can be repeated. We are now dealing with the operator N , eqn. (23), which

* I am indebted to my colleague Dr. S. Bochner for the proof of this.

by (18a) and (24) is an integral of the eqn. $L = 0$, and not with the original operator M . Thus we form

$$N^2 = N_{12}^2 + N_{23}^2 + N_{31}^2 \quad (42)$$

then the eqn. $N^2 L = L N^2$, analogous to (18a), also holds for N^2 ; so, if u is a solution of $L = 0$, $N^2 u$ is also a solution of the same equation, whence the following relation, analogous to (24), follows:

$$N^2 u = D u \quad (43)$$

D is a new "integration constant," which we shall soon find is $j(j+1) \left(\frac{h}{2\pi}\right)^2$, where j is half-integral and corresponds precisely to the total momentum of the motion used earlier. In the sense of § 9, p. 254, M^2 corresponds to our earlier l ; but, in contrast to N^2 , M^2 is not "constant," a fact we indicated on p. 96.

On account of (23) and after reduction, (42) runs

$$\begin{aligned} N^2 &= M^2 + \frac{h}{2\pi i} \sum \alpha_k \alpha_l M_{kl} + \frac{1}{4} \left(\frac{h}{2\pi i}\right)^2 \sum (\alpha_k \alpha_l)^2 \\ &= M^2 + \left(\frac{h}{2\pi i}\right)^2 \left(X - \frac{3}{4}\right) \end{aligned} \quad (43a)$$

Here we have used the equation of definition (34) of X and the equation that follows from the rules of commutation, namely

$$(\alpha_k \alpha_l)^2 = \alpha_k \alpha_l \alpha_k \alpha_l = -\alpha_l^2 \alpha_k^2 = -1 \quad (44)$$

Hence eqn. (43) assumes the form

$$\left\{ M^2 + \left(\frac{h}{2\pi i}\right)^2 \left(X - \frac{3}{4}\right) - D \right\} u = 0 \quad (45)$$

In X the products $\alpha_1 \alpha_2$, $\alpha_2 \alpha_3$, $\alpha_3 \alpha_1$ occur. We shall show that they can be allocated to our three quaternion quantities δ . It is evident that they correspond to them in their general properties; for, like the latter, they are anti-commutative in pairs and are normalised to -1 [cf. (44)]. The way in which this allocation is effected in detail is as follows:

Calculating with matrices of two rows means in the first place that we have to consider not *one* unknown function u but a *pair of functions* u_1 , u_2 , which take the place of the pair of variables x_1 , x_2 . Thus, by eqn. (40), $\delta_1 u$, for example, corresponding to the first or second row of this matrix, signifies

$$\delta_1 u = \begin{cases} i \cdot u_1 + 0 \cdot u_2 = i u_1, \\ 0 \cdot u_1 - i \cdot u_2 = -i u_2. \end{cases}$$

Linking up with (25) and (26a) we assume the two functions u_1 , u_2 to be of the form

$$u_1 = R_1 \Theta_1 e^{i m_1 \phi} e^{\frac{2\pi i}{h} E t}, \quad u_2 = R_2 \Theta_2 e^{i m_2 \phi} e^{\frac{2\pi i}{h} E t} \quad (46)$$

that is, with *different* m 's (as also with different R 's and Θ 's), but of course with a common proper value E , which is characteristic of the whole state, not of the individual variable of state. We insert our assumption (46) in (27). Since the constant C is to stand on the right-hand side, or, expressed more fully, the quantity Cu_1 in the first, Cu_2 in the second row of the equation, then in our assumption (46) giving the way in which the u 's depend on ϕ , the left-hand side of (27) may contain only the factors u_1 and u_2 in the first and second row. But then the matrix $\alpha_1\alpha_2$ that occurs in the left of (27) must be *diagonal* in form. The only one of the three quantities $\delta_1, \delta_2, \delta_3$, that has this property is δ_1 . Hence it follows that

$$\alpha_1\alpha_2 = \delta_1 = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix} \quad (47)$$

Eqn. (27) therefore resolves into the two following statements :

$$\left. \begin{aligned} \frac{h}{2\pi} \left(m_1 + \frac{1}{2} \right) u_1 &= Cu_1 \\ \frac{h}{2\pi} \left(m_2 - \frac{1}{2} \right) u_2 &= Cu_2 \end{aligned} \right\} \quad (48)$$

If we write $m_1 = m$, then it follows from (48) that

$$m_2 = m + 1, \quad C = \frac{h}{2\pi} (m + 1/2) \quad (48a)$$

Having set $\alpha_1\alpha_2$ equal to δ_1 we have left for the other two products $\alpha_2\alpha_3$ and $\alpha_3\alpha_1$ the quaternion quantities δ_2 and δ_3 . We choose

$$\alpha_2\alpha_3 = \delta_2 = \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix}, \quad \text{thus } \alpha_3\alpha_1 = \delta_2 = \begin{vmatrix} 0 & i \\ i & 0 \end{vmatrix} \quad (49)$$

and we shall soon show that the reverse choice is in contradiction to our assumption (46).

Our quantity X in (34) is then determined as follows :

$$\frac{h}{2\pi i} X = \begin{Bmatrix} iM_{12}u_1 - M_{23}u_2 + iM_{31}u_2, \\ -iM_{12}u_2 + M_{23}u_1 + iM_{31}u_1. \end{Bmatrix}$$

We here substitute from equation (23a) of this section and eqns. (41) and (41a) of § 9 and obtain

$$X = \begin{Bmatrix} i \frac{\partial u_1}{\partial \phi} + e^{-i\phi} \left(\cot \theta \frac{\partial u_2}{\partial \phi} + i \frac{\partial u_2}{\partial \theta} \right), \\ -i \frac{\partial u_2}{\partial \phi} - e + i\phi \left(\cot \theta \frac{\partial u_1}{\partial \phi} - i \frac{\partial u_1}{\partial \theta} \right) \end{Bmatrix} \quad (50)$$

The first row has all three terms multiplied by $e^{im\phi}$, the second by $e^{i(m+1)\phi}$, as follows from the form of u_1, u_2 , in eqn. (46), and the meaning of m_1, m_2 in eqn. (48a). Thus our choice (49) has stood the test; the reverse choice would lead to the individual terms of X depending differently on ϕ and is therefore to be rejected.

Like X so eqn. (45) is to be written in two rows. After the factor $e^{im\theta}$ or $e^{i(m+1)\theta}$, respectively, and the dependence on time have been cancelled, it runs

$$\left. \begin{aligned} A_1 R_1 \Theta_1 + i(m+1) \cot \theta R_2 \Theta_2 + i R_2 \frac{d\Theta_2}{d\theta} &= 0 \\ A_2 R_2 \Theta_2 - im \cot \theta R_1 \Theta_1 + i R_1 \frac{d\Theta_1}{d\theta} &= 0 \end{aligned} \right\} \quad (51)$$

with the temporary abbreviations

$$\begin{aligned} A_1 &= \left(\frac{2\pi i}{h}\right)^2 (M^2 - D) - m - \frac{3}{4} \\ A_2 &= \left(\frac{2\pi i}{h}\right)^2 (M^2 - D) + m + 1 - \frac{3}{4} \end{aligned}$$

Here the dependence on θ must also cancel out, if these equations are to hold simultaneously for all values of θ . This is the case if in the first and second equation the factors of iR_2 and iR_1 become proportional to Θ_1 and Θ_2 , respectively, and if, besides, the two factors A_1 and A_2 become independent of θ . The first condition demands that

$$\left. \begin{aligned} (m+1) \cot \theta \Theta_2 + \frac{d\Theta_2}{d\theta} &= a\Theta_1 \\ -m \cot \theta \Theta_1 + \frac{d\Theta_1}{d\theta} &= b\Theta_2 \end{aligned} \right\} \quad (52)$$

in which the two factors of proportionality a and b here introduced are constants. Elimination of Θ_2 from (52) gives for Θ_1 the differential equation

$$\frac{d^2\Theta_1}{d\theta^2} + \cot \theta \frac{d\Theta_1}{d\theta} + \left(m(m+1) - ab - \frac{m^2}{\sin^2 \theta}\right)\Theta_1 = 0.$$

By eqn. (1b) on p. 8 this is the differential equation of the spherical harmonic

$$\Theta_1 = P_l^m(\cos \theta) \quad (52a)$$

if we make

$$l_1(l_1 + 1) = m(m+1) - ab \quad (53a)$$

In the same way the elimination of Θ_1 from (52) gives

$$\frac{d^2\Theta_2}{d\theta^2} + \cot \theta \frac{d\Theta_2}{d\theta} + \left(m(m+1) - ab - \frac{(m+1)^2}{\sin^2 \theta}\right)\Theta_2 = 0,$$

that is, again, the differential equation of a spherical harmonic, namely of

$$\Theta_2 = P_{l_2}^{m+1}(\cos \theta) \quad (52b)$$

where now

$$l_2(l_2 + 1) = m(m+1) - ab \quad (53b)$$

The assumption (52a, b) of a spherical harmonic form for Θ_1 , Θ_2 , however, also satisfy the other condition that we made for allowing the eqn. (51) to be fulfilled, namely the constancy of the factors A_1 , A_2 . For M^2

denotes precisely the differential expression of the spherical harmonic functions [eqn. (43a) in § 9], so that on account of (52a, b) we get in the first and second eqn. (51), respectively (note the factor $\hbar/2\pi i$ in the present definition of M),

$$\left. \begin{aligned} \left(\frac{2\pi i}{\hbar}\right)^2 M^2 &= -l_1(l_1 + 1), \quad A_1 = -l_1(l_1 + 1) - D' - m - \frac{3}{4}, \\ \left(\frac{2\pi i}{\hbar}\right)^2 M^2 &= -l_2(l_2 + 1), \quad A_2 = -l_2(l_2 + 1) - D' + m + 1 - \frac{3}{4} \end{aligned} \right\} \quad (54)$$

We have here used the abbreviation

$$D' = \left(\frac{2\pi i}{\hbar}\right)^2 D. \quad (54a)$$

From (53a, b) it follows immediately that

$$l_2(l_2 + 1) = l_1(l_1 + 1)$$

and hence either $l_2 = l_1$ or $l_2 = -(l_1 + 1)$. But since the l 's, being proper values of the spherical harmonic problem, are necessarily positive we must exclude the second choice; so we may write

$$l_1 = l_2 = l. \quad (54b)$$

From (53a, b) we then also get

$$ab = m(m + 1) - l(l + 1). \quad (54c)$$

But, by (52), a and b are also determined individually, namely

$$b = -1, \quad a = l(l + 1) - m(m + 1) = (l - m)(l + m + 1) \quad (54d)$$

It then follows from (51) by eliminating the R_1 's and R_2 's and taking into account (52) that

$$0 = \begin{vmatrix} A_1 & ia \\ ib & A_2 \end{vmatrix} = A_1 A_2 + ab.$$

If we substitute for A_1 , A_2 , l_1 , l_2 and ab from (54) and (54 b, c), then a quadratic equation for D' results which, using the abbreviation,

$$\xi = l(l + 1) + \frac{3}{4} + D'$$

we may write

$$\xi^2 - \xi = l(l + 1) \text{ or } (\xi - \frac{1}{2})^2 = (l + \frac{1}{2})^2 \text{ or } \xi = \frac{1}{2} \pm (l + \frac{1}{2}).$$

Accordingly,

$$D' = \frac{1}{2} \pm (l + \frac{1}{2}) - l(l + 1) - \frac{3}{4} = \pm (l + \frac{1}{2}) - (l + \frac{1}{2})^2. \quad (55)$$

D' may also be written in the form

$$D' = \left\{ \begin{aligned} &-(l + \frac{1}{2})(l - \frac{1}{2}), \\ &-(l + \frac{3}{2})(l + \frac{1}{2}) \end{aligned} \right\}$$

corresponding to the positive or negative sign in (55). So if we set

$$l \mp \frac{1}{2} = j = \text{orbital moment of momentum} + \text{electron spin}$$

$$= \text{total moment of momentum}. \quad (56a)$$

then D' assumes the form

$$D' = -j(j + 1).$$

From this it follows by (54a) that

$$D = j(j+1)\left(\frac{h}{2\pi}\right)^2,$$

and by (43) that

$$N = \sqrt{j(j+1)}\frac{h}{2\pi}.$$

Thus we again find here the j^2 characteristically replaced by $j(j+1)$ and we also see how the quantum number j is not introduced by a special hypothesis but arises spontaneously out of the structure of the equation $L = 0$ and its integral N^2 .

To the two values of D' in (55) there correspond two values of our operator X . For if we apply eqn. (45) to one of the two solutions u_1 or u_2 , then we certainly get

$$\left(\frac{2\pi i}{h}\right)^2 M^2 u = -l(l+1)u$$

and eqn. (45) gives

$$[-l(l+1) + X - \frac{3}{4} - D']u = 0.$$

It therefore follows that

$$\begin{aligned} Xu &= [l(l+1) + \frac{3}{4} + D']u = [(l + \frac{1}{2})^2 + \frac{1}{4} + D']u \\ &= [\frac{1}{2} \pm (l + \frac{1}{2})]u. \end{aligned}$$

So we have for both functions u

$$\text{either } Xu = (l+1)u \text{ or } Xu = -lu. \quad (56)$$

This circumstance is the first indication that our pair of functions, that is, our two-dimensional scheme of matrices is not sufficient. The double value of X shows that there are not two but twice two functions (u_1, u_2, u_3, u_4) that are coupled with each other.

We next complete the representation of our functions u_1, u_2 so far used and then supplement them by the functions u_3, u_4 yet to be introduced. If we disregarded the time-factor that is common to all the u 's, we should have, according to (46), (48a), (52a, b) and (54b),

$$u_1 = R_1 P_l^m (\cos \theta) e^{im\phi}, \quad u_2 = R_2 P_l^{m+1} (\cos \theta) e^{i(m+1)\phi}.$$

Between R_1 and R_2 , however, there still exists the relationship expressed, for example, in the first row of (51). In view of (52) it states that

$$A_1 R_1 + ia R_2 = 0 \quad (57)$$

in which, on account of (54) and (55)

$$A_1 = -m - \frac{1}{2} \mp (l + \frac{1}{2}) \quad (57a)$$

From the law of areas our value D' in (55) followed, which, being a true integration constant, must be common to all solutions. Thus the double sign in (55) does not signify that there are two values of D' but rather that there are two different possibilities for the allocation of l to D' .

If according to definition we relate the upper sign in A_1, D' to u_1, u_2 , the lower * to u_3, u_4 , we have

$$R_2 = - \frac{i(m+l+1)}{a} R_1 = \frac{-i}{l-m} R_1$$

and as expressions for u_1, u_2 :

$$\left. \begin{aligned} u_1 &= R_1 P_l^m (\cos \theta) e^{im\phi} \\ u_2 &= \frac{-i}{l-m} R_1 P_l^{m+1} (\cos \theta) e^{i(m+1)\phi} \end{aligned} \right\} \quad (58)$$

The corresponding substitution for u_3, u_4 runs

$$u_3 = R_3 P_{l'}^{m'} (\cos \theta) e^{im'\phi}, \quad u_4 = R_4 P_{l'}^{m'+1} (\cos \theta) e^{i(m'+1)\phi} \quad (59)$$

In this we have already used the fact that the new numbers l', m' must agree with the consequences (48a) and (54b) deduced from the law of areas.

Having chosen the upper sign in (55) for l and u_1, u_2 , the lower sign belongs to l' and u_3, u_4 . So we conclude from (55) that

$$+ (l + \frac{1}{2}) - (l + \frac{1}{2})^2 = - (l' + \frac{1}{2}) - (l' + \frac{1}{2})^2$$

and this is equivalent to

$$(l' + 1)^2 = l^2.$$

On account of the positive character of l and l' it follows from this uniquely that

$$l' = l - 1 \quad (60)$$

There further results from (60) the restriction

$$l \neq 0 \quad (60a)$$

as otherwise l' would become negative.

From the allocation of the signs in (55) to u_1, u_2 and l on the one hand, and u_3, u_4 , and l' on the other, it further follows that in (56), too, the first value is to be allocated to u_1, u_2, l , the second to u_3, u_4, l' . Accordingly, eqn. (56) now runs

$$Xu = (l+1)u \text{ for } u = u_1, u_2 \quad (56a)$$

and

$$Xu = -l'u = (-l+1)u \text{ for } u = u_3, u_4 \quad (56b)$$

By means of the matrix of four rows

$$\epsilon = \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \quad (56c)$$

we compress the two cases into the single form

$$Xu = (1 + \epsilon l)u \quad (61)$$

* The fact that we may not take the same sign for u_1, u_2 and u_3, u_4 in A_1, D' (that is the upper or the lower sign for each of u_1, u_2, u_3, u_4 simultaneously) follows from the fact that otherwise, by (55), u_1, u_2, u_3, u_4 would acquire the same l -value (in the spherical harmonic function), and that would mean that $M^2 u = \text{const. } u$ would be true with the same constant for all u 's—which is impossible. For $M^2 u$ cannot be an integral of $L = 0$, since $(M^2 L - L M^2) u \neq 0$ (cf. sub-section C).

Like D' or D the quantity C introduced in (24) is also a true integration constant of the problem, which must have the same value for u_3, u_4 as for u_1, u_2 . Hence (48) shows directly that

$$m' = m.$$

To make our assumption (59) quite precise we must use eqn. (57) with R_3, R_4 in place of R_1, R_2 and with the lower sign of A_1 in (57a), l being replaced by $l' = l - 1$. In this way we obtain

$$R_4 = \frac{+i}{l+m} R_3,$$

and finally

$$\left. \begin{aligned} u_3 &= R_3 P_{l-1}^m (\cos \theta) e^{im\phi} \\ u_4 &= \frac{i}{l+m} R_3 P_{l-1}^{m+1} (\cos \theta) e^{i(m+1)\phi} \end{aligned} \right\} \quad (62)$$

But eqns. (58) and (62) do not exhaust all the possibilities of integration. In eqn. (57a) we could have carried the lower sign of A_1 over to u_1, u_2 . The values of u_1, u_2 then simply change places with those of u_3, u_4 ; in this process it is, of course, permissible to call the radial component of u_1 and u_3, R_1 and R_3 , respectively, as before. We shall write down the solutions that result in this way:

$$\left. \begin{aligned} u_1 &= R_1 P_{l-1}^m e^{im\phi} \\ u_2 &= \frac{i}{l+m} R_1 P_{l-1}^{m+1} e^{i(m+1)\phi} \end{aligned} \right\} \quad (58a)$$

$$\left. \begin{aligned} u_3 &= R_3 P_l^m e^{im\phi} \\ u_4 &= \frac{-i}{l-m} R_3 P_l^{m+1} e^{i(m+1)\phi} \end{aligned} \right\} \quad (62a)$$

Further, we then have

$$\begin{aligned} Xu &= (-l+1)u & \text{for } u = u_1, u_2, \\ Xu &= (l+1)u & \text{for } u = u_3, u_4, \\ \text{that is, } Xu &= (1-\epsilon l)u \end{aligned} \quad (61a)$$

with the same matrix ϵ as in eqn. (56c). Thus the two systems of solutions differ in X only in the sign of l , as we recognise by comparing (61a) with (61). For l the relation

$$l \neq 0,$$

again holds.

G. Determination of the Matrices α_k with Four Rows

We first show that α_k must be equal to our matrix ϵ of eqn. (56c). This follows from the fact that $\alpha_k (X-1)u$ is, with u , an integral of our equation $L=0$, or, expressed otherwise, that

$$\alpha_k (X-1)u = Fu \quad (63)$$

where F is an ordinary constant of integration (not a matrix). For, on account of (61), (61a), we see that (63) is equivalent to

$$\pm \alpha_i \epsilon u = F u \quad . \quad . \quad . \quad (64)$$

Hence it follows* that $F = \pm 1$ and $\alpha_i \epsilon = \pm 1$, and so, on account of $\alpha_i^2 = 1$:

$$\alpha_i = \pm \epsilon = \pm \begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{vmatrix} \quad . \quad . \quad . \quad (65)$$

It remains to prove that $\alpha_i(X - 1)$ is really an integral of the equation $L = 0$. The criterion for this is

$$\alpha_i(X - 1)L - L\alpha_i(X - 1) = 0 \quad . \quad . \quad . \quad (66)$$

To prove this we first consider the term with the factor -1 in (66), namely

$$-\alpha_i L + L\alpha_i \quad . \quad . \quad . \quad (66a)$$

Since α_i behaves anti-commutatively with respect to each of the three summands $\alpha_1 \Omega_1$, $\alpha_2 \Omega_2$, $\alpha_3 \Omega_3$, of L , but commutatively with respect to both the remaining $\alpha_4 \Omega_4$ and A , (66a) becomes equal to

$$2(\alpha_1 \alpha_i \Omega_1 + \alpha_2 \alpha_i \Omega_2 + \alpha_3 \alpha_i \Omega_3) \quad . \quad . \quad . \quad (66b)$$

We next consider one of the three terms

$$\alpha_k \alpha_i M_{kl}$$

of which X is made up. Using the cyclically generalised eqn. (14) we transform

$$M_{kl}L \text{ into } LM_{kl} + \alpha_k \Omega_k - \alpha_l \Omega_l,$$

and obtain as the contribution of this term to (66)

$$\{ \alpha_i \alpha_k \alpha_l L - L \alpha_i \alpha_k \alpha_l \} M_{kl} + \alpha_i \alpha_k \alpha_l (\alpha_k \Omega_k - \alpha_l \Omega_l) \quad . \quad (66c)$$

Here the second summand is equal to

$$\alpha_i \alpha_k \Omega_k + \alpha_i \alpha_l \Omega_l,$$

and summed up over all three terms of X , this gives

$$2 \sum_1^3 \alpha_i \alpha_j \Omega_j = -2 \sum_1^3 \alpha_j \alpha_i \Omega_j \quad . \quad . \quad . \quad (66d)$$

So it cancels with (66b). Lastly we show that the first summand in (66c) also vanishes for all terms of L with the exception of that whose index—say j —differs from k , l and 4. For this term gives

$$2\alpha_i \alpha_k \alpha_j \Omega_j M_{kl}$$

* The two signs in $F = \pm 1$ are here not uniquely allocated to the two systems of solutions (53), (63) and (58a), (62a); rather, $F = \pm 1$ can belong to either system, and hence also $\alpha_i = \pm \epsilon$.

If we sum it up over all three pairs of indices $k, l = 1, 2, j = 3; k, l = 2, 3, j = 1$; and $k, l = 3, 1, j = 2$ we get

$$2\alpha_4(\alpha_1\alpha_2\alpha_3\Omega_3M_{12} + \alpha_2\alpha_3\alpha_1\Omega_1M_{23} + \alpha_3\alpha_1\alpha_2\Omega_2M_{31}) = 2\alpha_4\alpha_1\alpha_2\alpha_3(\Omega_3M_{12} + \Omega_1M_{23} + \Omega_2M_{31})$$

and it can easily be verified that this last contribution vanishes, because $\Omega_3M_{12} + \Omega_1M_{23} + \Omega_2M_{31}$ becomes equal to zero. Hence eqn. (66) is proved.

Having found in (65) the meaning of (α_4) we can now also determine $\alpha_1, \alpha_2, \alpha_3$. The relation

$$\alpha_k\epsilon = -\epsilon\alpha_k, \quad k = 1, 2, 3,$$

which follows directly from our postulate (4) serves as our starting-point.

If we assume in general that

$$\alpha_k = \begin{vmatrix} A_{11} & A_{12} & . & A_{14} \\ A_{21} & A_{22} & . & . \\ . & . & . & . \\ A_{41} & . & . & A_{44} \end{vmatrix},$$

then we get on account of (56c) [by the rule rows multiplied by columns of p. 267, applied to our four-row matrices],

$$\alpha_k\epsilon = \begin{vmatrix} A_{11} & A_{12} & -A_{13} & -A_{14} \\ A_{21} & A_{22} & -A_{23} & -A_{24} \\ A_{31} & A_{32} & -A_{33} & -A_{34} \\ A_{41} & A_{42} & -A_{43} & -A_{44} \end{vmatrix}$$

$$-\epsilon\alpha_k = \begin{vmatrix} -A_{11} & -A_{12} & -A_{13} & -A_{14} \\ -A_{21} & -A_{22} & -A_{23} & -A_{24} \\ A_{31} & A_{32} & A_{33} & A_{34} \\ A_{41} & A_{42} & A_{43} & A_{44} \end{vmatrix}$$

From this it follows that in the scheme of all three α_k 's only the right upper and left lower half-squares are occupied, the other half-squares are zeros. The α_k 's are therefore once again reducible to matrices with two rows, which, for α_1 , we shall call a and a' , for α_2 , b and b' , for α_3 , c and c' :

$$\alpha_1 = \begin{vmatrix} 0 & a \\ a' & 0 \end{vmatrix} \quad \alpha_2 = \begin{vmatrix} 0 & b \\ b' & 0 \end{vmatrix} \quad \alpha_3 = \begin{vmatrix} 0 & c \\ c' & 0 \end{vmatrix} \quad (67)$$

Besides the a, \dots, c' , the zeros here also denote schemes of two rows.

$$\text{The condition} \quad 1 = \alpha_1^2 = \begin{vmatrix} aa' & 0 \\ 0 & a'a \end{vmatrix}$$

here denotes that we must have

$$aa' = a'a = 1 \quad (= \text{diagonal matrix}).$$

Likewise we must have

$$bb' = b'b = 1, \quad cc' = c'c = 1 \quad \dots \quad (68a)$$

We finally form $\alpha_i \alpha_l$ by means of (67) and compare the result of the multiplication with the expression (47), (49) in section F:

$$\alpha_1 \alpha_2 = \begin{vmatrix} \delta_1 & 0 \\ 0 & \delta_1 \end{vmatrix}, \quad \alpha_2 \alpha_3 = \begin{vmatrix} \delta_3 & 0 \\ 0 & \delta_3 \end{vmatrix}, \quad \alpha_3 \alpha_1 = \begin{vmatrix} \delta_3 & 0 \\ 0 & \delta_3 \end{vmatrix}. \quad (69)$$

Written in this form with four rows (the δ 's as well as the zeros are matrices with two rows) this expression states, in agreement with our conclusions in F, that it holds for the pair of variables u_1, u_2 just as much as for u_3, u_4 . On the other hand, it follows from (67) that

$$\alpha_1 \alpha_2 = \begin{vmatrix} ab' & 0 \\ 0 & a'b \end{vmatrix}, \quad \alpha_2 \alpha_3 = \begin{vmatrix} bc' & 0 \\ 0 & b'c \end{vmatrix}. \quad (69a)$$

A comparison of the two values of $\alpha_1 \alpha_2$ gives

$$ab' = a'b = \delta_1.$$

Multiplication by a and a' , respectively, gives, by (68),

$$b = a\delta_1, \quad b' = a'\delta_1. \quad (69b)$$

and multiplying according to (68a), having regard to the normalisation $\delta^2 = -1$ we have

$$1 = a\delta_1 a' \delta_1, \quad \delta_1 = -a\delta_1 a' \\ \delta_1 a = -a\delta_1. \quad (70)$$

At the same time we have

$$\delta_1 a' = -a'\delta_1, \quad \delta_1 b = -b\delta_1, \quad \delta_1 b' = -b'\delta_1. \quad (70a)$$

But, taking (40) into account, we calculate that

$$\delta_1 a = \begin{vmatrix} i & 0 \\ 0 & -i \end{vmatrix} \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = \begin{vmatrix} ia_{11} & ia_{12} \\ -ia_{21} & -ia_{22} \end{vmatrix}, \\ -a\delta_1 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \begin{vmatrix} -i & 0 \\ 0 & i \end{vmatrix} = \begin{vmatrix} -ia_{11} & ia_{12} \\ -ia_{21} & ia_{22} \end{vmatrix}.$$

Hence, on account of (70),

$$a_{11} = a_{22} = 0. \quad (71)$$

Likewise from (70a)

$$a_{11}' = a_{22}' = b_{11} = b_{22} = b_{11}' = b_{22}' = 0. \quad (71a)$$

Thus at present two elements remain in each of the matrices a, a', b, b' ; but by (68) these are alternately reciprocal,

$$a_{12}a_{21}' = 1, \quad a_{21}a_{12}' = 1. \quad (72)$$

Likewise, by (68a),

$$b_{12}b_{21}' = 1, \quad b_{21}b_{12}' = 1. \quad (72a)$$

and by (69b)

$$b_{12} = -ia_{12}, \quad b_{21} = ia_{21}, \quad b_{12}' = -ia_{12}', \quad b_{21}' = ia_{21}'.$$

Recapitulating, we have, if we set $a_{12} = p, a_{21} = q$

$$\left. \begin{aligned} a &= \begin{vmatrix} 0 & p \\ q & 0 \end{vmatrix}, & a' &= \begin{vmatrix} 0 & 1/q \\ 1/p & 0 \end{vmatrix}, \\ b &= \begin{vmatrix} 0 & -ip \\ iq & 0 \end{vmatrix}, & b' &= \begin{vmatrix} 0 & -i/q \\ i/p & 0 \end{vmatrix} \end{aligned} \right\} \quad (73)$$

So far we have made use of only the first of the relations (69). The second gives $bc' = b'c = \delta_3$ and hence, on account of (68a) and (40)

$$\left. \begin{aligned} c = b\delta_3 &= \begin{vmatrix} 0 & -ip \\ iq & 0 \end{vmatrix} \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = \begin{vmatrix} -ip & 0 \\ 0 & -iq \end{vmatrix} \\ c' = b'\delta_3 &= \begin{vmatrix} 0 & -i/q \\ i/p & 0 \end{vmatrix} \begin{vmatrix} 0 & -1 \\ 1 & 0 \end{vmatrix} = \begin{vmatrix} -i/q & 0 \\ 0 & -i/p \end{vmatrix} \end{aligned} \right\} \quad (73a)$$

But (68a) also makes

$$cc' = \begin{vmatrix} -p/q & 0 \\ 0 & -q/p \end{vmatrix} = 1$$

and this means that

$$q = -p \quad (74)$$

The value of p remains undetermined, for we see immediately that the following expressions for the α 's, which result from (67), (73) and (73a)

$$\left. \begin{aligned} \alpha_1 &= \begin{vmatrix} 0 & 0 & 0 & p \\ 0 & 0 & -p & 0 \\ 0 & -1/p & 0 & 0 \\ 1/p & 0 & 0 & 0 \end{vmatrix}, & \alpha_2 &= \begin{vmatrix} 0 & 0 & 0 & -ip \\ 0 & 0 & -ip & 0 \\ 0 & i/p & 0 & 0 \\ i/p & 0 & 0 & 0 \end{vmatrix} \\ \alpha_3 &= \begin{vmatrix} 0 & 0 & -ip & 0 \\ 0 & 0 & 0 & ip \\ i/p & 0 & 0 & 0 \\ 0 & -i/p & 0 & 0 \end{vmatrix} \end{aligned} \right\} \quad (75)$$

together with our expression for α_4 satisfy the initial postulates (3) and (4) for the α 's for arbitrary values of p . Accordingly it is not necessary to treat the further conditions contained in the eqns. (69), as these must be fulfilled of themselves.

In the next section we shall require besides the α 's also the products $\alpha\epsilon$:

$$\left. \begin{aligned} \alpha_1\epsilon &= \begin{vmatrix} 0 & 0 & 0 & -p \\ 0 & 0 & p & 0 \\ 0 & -1/p & 0 & 0 \\ 1/p & 0 & 0 & 0 \end{vmatrix}, & \alpha_2\epsilon &= \begin{vmatrix} 0 & 0 & 0 & ip \\ 0 & 0 & ip & 0 \\ 0 & i/p & 0 & 0 \\ i/p & 0 & 0 & 0 \end{vmatrix} \\ \alpha_3\epsilon &= \begin{vmatrix} 0 & 0 & ip & 0 \\ 0 & 0 & 0 & -ip \\ i/p & 0 & 0 & 0 \\ 0 & -i/p & 0 & 0 \end{vmatrix} \end{aligned} \right\} \quad (76)$$

Here we should apply a remark about the behaviour of the α 's in Lorentz-Einstein transformations. By their definitions as matrices with four rows the α 's transform in an orthogonal substitution like four-vectors. The Ω 's do the same, by p. 102. But the exact proof of the invariance of $\Sigma \alpha_i \Omega_i$ in eqn. (5) requires further consideration, which we must omit here.

H. The Radial Differential Equations of the Kepler Problem

We revert to eqn. (36), first dealing with the equations that belong to the system of solutions (58), (61) and (62), which we shall call I for short; we shall then call eqns. (58a), (61a) and (62a) system II.

Substituting for γ from (29) and for X from (61), we get for eqn. (36)

$$\sum_1^3 \alpha_k \frac{x_k}{r} \left(\frac{\partial}{\partial r} + \frac{1 + \epsilon^2}{r} \right) u = - \frac{2\pi}{\hbar c} \{ \alpha_4 (E - V) + E_0 \} u \quad (77)$$

For x_k we write

$$x_1 = r \sin \theta \cos \phi, \quad x_2 = r \sin \theta \sin \phi, \quad x_3 = r \cos \theta.$$

We get α_k and $\alpha_{k\epsilon}$ out of (75) and (76), from the first row of the corresponding matrices. So we obtain as the first term of (77)

$$p \sin \theta \cos \phi \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) u_4 - ip \sin \theta \sin \phi \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) u_4 \\ - ip \cos \theta \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) u_3$$

or, with terms conveniently collected,

$$p \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) \{ \sin \theta e^{-i\phi} u_4 - i \cos \theta u_3 \} \quad (77, 1)$$

We obtain corresponding results for the subsequent rows:

$$- p \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) \{ \sin \theta e^{+i\phi} u_3 - i \cos \theta u_4 \} \quad (77, 2)$$

$$- \frac{1}{p} \left(\frac{\partial}{\partial r} + \frac{1 + l}{r} \right) \{ \sin \theta e^{-i\phi} u_2 - i \cos \theta u_1 \} \quad (77, 3)$$

$$\frac{1}{p} \left(\frac{\partial}{\partial r} + \frac{1 + l}{r} \right) \{ \sin \theta e^{+i\phi} u_1 - i \cos \theta u_2 \} \quad (77, 4)$$

We write down the first of the equations bracketed in (77) after substituting u_4 and u_3 from (62), and u_1 from (58) in (77, 1), and taking (65) into consideration,

$$\left. \begin{aligned} p \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) \{ \sin \theta e^{-i\phi} R_4 e^{i(m+1)\phi} P_{l-1}^{m+1}(\cos \theta) \\ - i \cos \theta R_3 e^{im\phi} P_{l-1}^m(\cos \theta) \} \\ = \frac{2\pi}{\hbar c} [\mp (E - V) - E_0] R_1 e^{im\phi} P_l^m(\cos \theta) \end{aligned} \right\} \quad (78)$$

We see that the dependence on ϕ drops out; the same is the case* for θ . Hence eqn. (78) may be written, as an easy calculation shows,

$$ip \left(\frac{\partial}{\partial r} + \frac{1 - l}{r} \right) R_4 = \frac{2\pi}{\hbar c} [\mp (E - V) - E_0] R_2 \quad (79, 1)$$

* To prove this, use the well-known formulae

$$P_l' = x P_{l-1}' + l P_{l-1}, \quad P_{l-1}' = x P_l' - l P_l;$$

by successive differentiation we get corresponding formulae for the associated functions. It is convenient to transform from R_3 and R_1 to R_4 and R_2 by means of the relations given in (58) and (62).

The second row of (77) gives exactly the same result. We shall not prove this, as it is not difficult and is of no interest. For the third row we get

$$\frac{1}{ip} \left(\frac{\partial}{\partial r} + \frac{1+l}{r} \right) R_2 = \frac{2\pi}{hc} [\pm (E - V) - E_0] R_4 \quad (79, 3)$$

The fourth row again gives the same result as the third.

If we now consider the case of the system II, in (36) only X changes, and in X we simply replace $+l$ by $-l$, according to the remark on eqn. (61a). Accordingly eqns. 77, 1 to 4, remain unaltered except for the substitution of $\mp l$ by $\pm l$ respectively in the operator

$$\frac{\partial}{\partial r} + \frac{1 \mp l}{r};$$

for the calculation of the matrices $\alpha_1, \alpha_2, \alpha_3$ is quite independent of the choice of the systems I or II.

We then get for the first and second row of eqn. (77)

$$ip \left(\frac{\partial}{\partial r} + \frac{1+l}{r} \right) R_1 = \frac{2\pi}{hc} [\mp (E - V) - E_0] R_2 \quad (80, 1)$$

and for the third and fourth row

$$\frac{1}{ip} \left(\frac{\partial}{\partial r} + \frac{1-l}{r} \right) R_2 = \frac{2\pi}{hc} [\pm (E - V) - E_0] R_4 \quad (80, 3)$$

We see at once that the eqns. (79, 1, 3) transform into the eqns. (80, 1, 3), if we replace l by $-l$ and conversely. We may combine the two pairs of equations into one by setting,*

$$k = \pm l, \quad k = \pm 1, \pm 2, \pm 3, \dots$$

For then we have

$$\left. \begin{aligned} ip \left(\frac{\partial}{\partial r} + \frac{1-k}{r} \right) R_1 &= \frac{2\pi}{hc} [\mp (E - V) - E_0] R_2; \\ \frac{1}{ip} \left(\frac{\partial}{\partial r} + \frac{1+k}{r} \right) R_2 &= \frac{2\pi}{hc} [\pm (E - V) - E_0] R_4 \end{aligned} \right\} \quad (81)$$

For ip we introduce the abbreviation

$$ip = g. \quad (82)$$

J. The Formula for the Fine-Structure

The first step towards integration consists, as always, in investigating the asymptotic behaviour. So, in (81), we strike out all terms with

* The symbol k is to call to memory the quantum number k of older theory and, in common with it, has the property that $k = 0$ must be excluded, cf. eqn. (60a). In our present nomenclature the S-term is given by $k = 1$ (cf. p. 286) just as in the older quantum theory. But our present k differs from our earlier k in being able to assume the values of negative integers besides positive integers. Dirac uses in place of k the letter j (likewise with the double sign). This letter does not seem to us happily chosen as j must be reserved later for the inner quantum number, which cannot be dispensed with. We have already used j in this sense on pp. 269 and 272. Our nomenclature agrees with that of O. G. Darwin, Proc. Roy. Soc., 112, 654 (1926).

$1/r$, which also include V , and obtain

$$\begin{aligned}\frac{1}{g} \frac{dR_2}{dr} &= \frac{2\pi}{\hbar c} (\pm E - E_0) R_4, \\ g \frac{dR_4}{dr} &= \frac{2\pi}{\hbar c} (\mp E - E_0) R_2.\end{aligned}$$

We here make the assumption, which is characteristic for linear differential equations with constant coefficients

$$R_2 = A_1 e^{-\lambda r}, \quad R_4 = A_2 e^{-\lambda r} \quad (83)$$

and we find as the equations of condition

$$\begin{aligned}-\frac{\lambda}{g} A_1 &= \frac{2\pi}{\hbar c} (\pm E - E_0) A_2, \\ -g \lambda A_2 &= \frac{2\pi}{\hbar c} (\mp E - E_0) A_1.\end{aligned}$$

From this it follows, with the assumption $E < E_0$ (line-spectrum) and with the choice of the correct (positive) sign of the root

$$\lambda = + \frac{2\pi}{\hbar c} \sqrt{E_0^2 - E^2} \quad (83a)$$

As earlier in the Kepler problem we use as the independent variable

$$\rho = 2\lambda r$$

and supplement the asymptotic assumption (83) to an exact expression by introducing as independent variables two functions v_1 and v_2 of ρ (dashes denote differential quotients with respect to ρ):

$$\left. \begin{aligned}R_2 &= v_1 e^{-\rho/2}, \quad R_4 = v_2 e^{-\rho/2} \\ \frac{dR_2}{d\rho} &= 2\lambda \left(v_1' - \frac{1}{2} v_1 \right) e^{-\rho/2} \\ \frac{dR_4}{d\rho} &= 2\lambda \left(v_2' - \frac{1}{2} v_2 \right) e^{-\rho/2}\end{aligned} \right\} \quad (84)$$

Eqs. (81), after division by 2λ and omission of the factor $e^{-\rho/2}$, then give

$$\left. \begin{aligned}\frac{1}{g} \left[v_1' + \left(\frac{1+k}{\rho} - \frac{1}{2} \right) v_1 \right] &= \left(\pm \frac{\alpha Z}{\rho} + \frac{1}{2} \frac{\pm E - E_0}{\sqrt{E_0^2 - E^2}} \right) v_2 \\ g \left[v_2' + \left(\frac{1-k}{\rho} - \frac{1}{2} \right) v_2 \right] &= \left(\mp \frac{\alpha Z}{\rho} + \frac{1}{2} \frac{\mp E - E_0}{\sqrt{E_0^2 - E^2}} \right) v_1\end{aligned} \right\} \quad (85)$$

For V we have substituted the expression $-Ze^2/r$ which is characteristic of the Kepler problem, so that

$$\frac{2\pi}{\hbar c} V = - \frac{2\pi e^2 Z}{\hbar c r} = - \frac{\alpha Z}{r}$$

(α = the fine-structure constant).

Passing on to investigate the origin we set

$$v_1 = \rho^\gamma \sum a_\gamma \rho^\gamma, \quad v_2 = \rho^\gamma \sum b_\gamma \rho^\gamma$$

with the same exponent γ but with different coefficients a_v, b_v . γ is determined by comparing in (85) the coefficient of $\rho^{\gamma-1}$ on both sides:

$$\frac{1}{g}(\gamma + 1 + k)a_0 = \pm \alpha Z b_0, \quad g(\gamma + 1 - k)b_0 = \mp \alpha Z a_0.$$

By multiplying these equations together term by term a_0, b_0 drop out and there is left

$$(\gamma + 1)^2 - k^2 = -\alpha^2 Z^2,$$

from which, if the sign of the root is correctly chosen (positive), we get*

$$\gamma = -1 + \sqrt{k^2 - \alpha^2 Z^2}. \quad (86)$$

We next calculate the recurrence formula for the a_v 's, b_v 's by equating the coefficients of the power $\rho^{\gamma+v}$ on both sides of eqns. (85).

We temporarily introduce the abbreviations

$$f_1 = \frac{\pm E - E_0}{\sqrt{E_0^2 - E^2}}, \quad f_2 = \frac{\mp E - E_0}{\sqrt{E_0^2 - E^2}} \quad (87)$$

and take the terms with a_{v+1}, b_{v+1} over to the left, those with a_v, b_v over to the right:

$$\left. \begin{aligned} (v + 1 + \gamma + 1 + k) \frac{a_{v+1}}{g} \mp \alpha Z b_{v+1} &= \frac{1}{2} \left(\frac{a_v}{g} + f_1 b_v \right) \\ (v + 1 + \gamma + 1 - k) b_{v+1} g \pm \alpha Z a_{v+1} &= \frac{1}{2} (b_v g + f_2 a_v) \end{aligned} \right\} \quad (88)$$

We see at once that the determinant of the a_v 's, b_v 's on the right-hand side is equal to zero:

$$\begin{vmatrix} 1/g & f_1 \\ f_2 & g \end{vmatrix} = 1 - 1 = 0, \quad \text{on account of } f_1 f_2 = 1,$$

whereas the determinant of the a_{v+1} 's, b_{v+1} 's is *not* equal to zero. If, then, we determine the a_v 's, b_v 's so that the right-hand sides become equal to zero, then the a_{v+1} 's, b_{v+1} 's necessarily vanish, and likewise all the subsequent coefficients. Our recurrence formula is broken off at a certain point. *The condition for this break (polynomial condition)* is, therefore,

$$\frac{a_v}{b_v} = -g f_1 \quad (89)$$

To determine the ratio a_v/b_v in a second way we proceed as follows. In (88) we replace $v + 1$ by v and multiply the first of the resulting

* It is to be observed that γ becomes negative for $k = \pm 1$ (the S-term or, of, the table on p. 288, the $P_{1/2}$ term), although only very slightly so if Z is not too great, namely of the order $-\alpha^2 Z^2/2$. From this it follows that the radial component R of the proper functions for $n = 0$ becomes infinite, contrary to our "boundary condition," which requires the proper functions to be continuous without exception. We have already made the same remark in § 9, Chap. I, p. 115. But we there hinted in a footnote at a wider formulation of the boundary conditions which would remove this difficulty. It runs: instead of requiring the proper functions to be continuous without exception, we postulate that it is to be integrable quadratically, which is already demanded in the condition of normalisation $\int \psi \psi^* d\tau = 1$. In our case this more comprehensive condition is obviously fulfilled for all values that actually occur in the periodic system.

equations by g , the second by f_1 . The right-hand sides then become equal, and so also the left-hand sides. We have

$$(\nu + \gamma + 1 + k)a_\nu \mp \alpha Z b_\nu g = g f_1(\nu + \gamma + 1 - k)b_\nu \pm f_1 \alpha Z a_\nu.$$

From this it follows that

$$\frac{a_\nu}{b_\nu} = \frac{g f_1(\nu + \gamma + 1 - k) \pm \alpha Z g}{\nu + \gamma + 1 + k \mp f_1 \alpha Z} \quad (90)$$

This holds for every index ν . If the recurrence is to cease for a given ν , say for $\nu = n_r$, then by (89) and (90) we must set

$$-(n_r + \gamma + 1 - k) \mp \frac{\alpha Z}{f_1} = n_r + \gamma + 1 + k \mp f_1 \alpha Z,$$

and hence, with the value (86) for γ , we have, collecting terms, that

$$n_r + \sqrt{k^2 - \alpha^2 Z^2} = \pm \frac{\alpha Z}{2} \left(f_1 - \frac{1}{f_1} \right) \quad (91)$$

n_r is the common degree of the polynomial components of v_1 and v_2 that result in this way. According to the meaning of f_1 we have

$$\pm \frac{1}{2} \left(f_1 - \frac{1}{f_1} \right) = \frac{E}{\sqrt{E_0^2 - E^2}} = \left\{ \left(\frac{E_0}{E} \right)^2 - 1 \right\}^{-\frac{1}{2}}.$$

Thus eqn. (91) states that

$$\left\{ \begin{aligned} \left(\frac{E_0}{E} \right)^2 - 1 &= \frac{\alpha^2 Z^2}{(n_r + \sqrt{k^2 - \alpha^2 Z^2})^2} \\ \frac{E}{E_0} &= \left\{ 1 + \frac{\alpha^2 Z^2}{(n_r + \sqrt{k^2 - \alpha^2 Z^2})^2} \right\}^{-\frac{1}{2}} \end{aligned} \right\} \quad (92)$$

This is the fine-structure formula (27) of I, p. 417, with slight differences of nomenclature. It has been derived simultaneously by W. Gordon and C. G. Darwin† from Dirac's theory of the electron.*

The inconsistencies in the wave-mechanical treatment in § 9, Chap. I, have now been removed. Compared with the older theory in I, Chap. 6, this shows an advance in that *every term-level is counted twice*, on account of the double sign of k , *except the highest level, for each given principal quantum number*, just as Fig. 11 demands.

¶ We prove this as follows. The principal quantum number n is now determined as a quantum sum by

$$n = n_r + |k| \quad (93)$$

corresponding to the denominator of the term formula (92) for $\alpha = 0$, in which the root sign, according to the definition of γ in (86), is to be taken as positive and hence becomes $|k|$ for $\alpha = 0$. The term level of highest energy is that for which $n_r = 0$; it corresponds to the circular orbits of the older theory. By (93) $|k|$ would be n , that is $k = \pm n$ for this level. We easily convince ourselves that only $k = +n$ is permissible, namely when we choose $\alpha_+ = +\epsilon$ [see eqn. (65)], or else only $k = -n$, if α_- is to equal $-\epsilon$.

* Zeitschr. f. Phys., 48, 11 (1928).

† Cf. the reference in Note *, p. 288.

This has its origin in the particular position occupied by the case $n_r = 0$ in the recurrence formula. For $n_r = 0$ the components of the polynomial in v_1, v_2 become of zero degree, that is equal to a_0 or b_0 , respectively. Instead of the recurrence-formula (88) we then have to satisfy the two equations

$$\begin{cases} (\gamma + 1 + k)a_0/g \mp \alpha Z b_0 = 0 \\ (\gamma + 1 - k)b_0/g \pm \alpha Z a_0 = 0 \end{cases} \quad (94)$$

for a_0, b_0 .

In addition we have for $v = n_r = 0$, by (89),

$$\frac{a_0}{b_0} = -gf_1 \quad (95)$$

Our determination of γ in (86) ensured the vanishing of the determinant of the a_0 's, b_0 's in (94). So we get from (94) and (95)

$$\frac{a_0}{b_0} = \pm \frac{\alpha Z g}{\gamma + 1 + k} = -gf_1,$$

and so, taking into account (86) and (87)

$$\frac{\alpha Z}{\sqrt{k^2 - \alpha^2 Z^2} + k} = \frac{-E \pm E_0}{\sqrt{E_0^2 - E^2}} = \begin{cases} \sqrt{\frac{E_0 - E}{E_0 + E}} \\ -\sqrt{\frac{E_0 + E}{E_0 - E}} \end{cases} \quad (96)$$

If we now choose $\alpha_4 = +\epsilon$, then the *upper*, the *positive* of the two preceding values applies. But then the left-hand is positive only if $k > 0$. Hence it necessarily follows from this that $k = +n$. For $\alpha_4 = -\epsilon$ it would in the same way follow that $k < 0$, that is $k = -n$. In each case the highest level ($n_r = 0$, circular orbit) is simple, as was asserted.

We shall now show how in definite instances the quantum numbers are to be allocated to the different levels. Let us choose $n = 3$. By (93) $|k|$ then has the values 1, 2, 3, whereas at the same time n_r becomes equal to 2, 1, 0. From the definition of k we have $k = \pm l$ with $l = 1, 2, 3, \dots$. But this l is by no means identical with the "azimuthal quantum" l of Schrödinger's theory (§ 7, Chap. I). To distinguish between these two numbers we shall now write h for the l of this section.

We find the connexion between h and l by going to the limit $c = \infty$ in the differential equations (85) of our problem. These solutions must then agree with those of Schrödinger's non-relativistic equation. We set $\alpha_4 = +\epsilon$; then the highest level belongs to $k = n$, the lower levels to $k = \pm(n-1), \pm(n-2), \dots$. On account of $\alpha_4 = +\epsilon$ the upper sign is to be used in all the equations of this and the preceding section. If we let $c \rightarrow \infty$, we must get from (81) Schrödinger's differential equation of the Kepler problem. By eliminating R_2 from eqn. (81) we arrive at the relation*

$$\left\{ \frac{\partial^2}{\partial r^2} + \frac{2}{r} \frac{\partial}{\partial r} - \frac{k(k-1)}{r^2} \mp \frac{8\pi^2 m}{h^2} (W - V) \right\} R_1 = 0 \quad (97)$$

* R_2 vanishes for $c \rightarrow \infty$, as is easily seen by considering eqn. (81).

where E has been set equal to $W + E_0$ [cf. also eqn. (40a), p. 112]. It agrees except for unessential differences in nomenclature with eqn. (3) of p. 60, if we make

$$l(l+1) = k(k-1) \quad (98)$$

Consequently $l = k - 1$ or $= -k$, and on account of the positive character of l we have for $k > 0$

$$l = k - 1 = h - 1 \quad (99a)$$

and for $k < 0$

$$l = -k = |k| = h \quad (99b)$$

The corresponding j -values (inner quantum numbers) come out according to eqn. (55a) of p. 272 for $k > 0$, that is, for the system I as

$$j = l + \frac{1}{2} = k - 1 + \frac{1}{2} \quad (100a)$$

We have here to observe that eqns. (97) and (98) refer to R_+ and so to u_3, u_4 and hence to the lower sign of eqn. (55a). Thus j becomes equal to $\frac{3}{2}, \frac{3}{2}, \frac{1}{2}$ for $k = 3, 2, 1$. On the other hand, $k < 0$ gives

$$j = l - \frac{1}{2} = |k| - \frac{1}{2} \quad (100b)$$

since for $k < 0$ system II holds, and we must reverse the sign in (55a) [cf. the arguments on pp. 274 and 275], and then again use the sign corresponding to u_3, u_4 in (55a). Thus we have $j = \frac{3}{2}, \frac{1}{2}$ for $k = -2, -1$.

The complete allocation of quantum numbers and levels is given in our case, $n = 3$, by the following scheme :

n .	n_r	$ k $.	k .	l .	Term.	j .
3	0	3	3	2	D	$\frac{5}{2}$
	1	2	$\begin{Bmatrix} -2 \\ +2 \end{Bmatrix}$	$\begin{Bmatrix} 2 \\ 1 \end{Bmatrix}$	P	$\begin{Bmatrix} \frac{3}{2} \\ \frac{5}{2} \end{Bmatrix}$
	2	1	$\begin{Bmatrix} -1 \\ +1 \end{Bmatrix}$	$\begin{Bmatrix} 1 \\ 0 \end{Bmatrix}$	S	$\begin{Bmatrix} \frac{1}{2} \\ \frac{3}{2} \end{Bmatrix}$

The levels with *different* $|k|$'s form relativistic doublets.

The generally double nature of the energy-levels having the *same* $|k|$ (screening doublets) but different *signs* for k , on the other hand, remains latent in the pure Coulomb field. But it asserts itself at once if a magnetic field or an inner atomic central field becomes added, that is, in the Zeeman effect or in the case of atoms which are not of the hydrogen type. This gives the *alkali-like* character of hydrogen, as was postulated in Fig. 11 by the new notation (on the right). The consequences of this notation for the selection rules discussed earlier may now be derived quantitatively from the representation of $u_1 \dots u_4$ in spherical harmonics. In the same way the anomalous Zeeman effect in the case of hydrogen and the alkalis may be cogently treated, and also the system of terms for atoms in general may be placed on a quantitative basis.

Dirac's theory in conjunction with Pauli's Exclusion Principle has inspired Eddington to go still further and has led him to draw an unusually bold inference. He counts up the possibilities according to which

the Dirac matrices with four rows can be formed and combined and finds from them the following value for the reciprocal of the fine-structure constant:—

$$\frac{1}{\alpha} = 16 + \frac{16.15}{2} = 136.$$

This would bring the electronic charge e into a simple numerical relationship with the fundamental constants c and h , and would signify that the existence of the elementary quantum of electricity could be conceived as a consequence of the theory of relativity and the quantum theory.

Unfortunately experimental evidence is not at the present time in favour of Eddington's theory. The older measurements of Millikan give, as we know, the value 137 rather than 136 for $1/\alpha$. The most recent measurements by A. H. Compton and J. A. Bearden give the value $1/\alpha = 136.6$. A final decision on the experimental side of the question is just as little possible at present as on the theoretical, that is, whether Eddington's conjecture is justifiable. At any rate this also indicates what extraordinary importance Dirac's theory has at present and will have in future for dealing with the fundamental questions of natural philosophy,

NOTES

1. On the Continuous Proper Functions in the Hydrogen Spectrum, their Asymptotic Representation and their Normalisation

WE have represented the radial part R of the *discrete* proper functions in the hydrogen spectrum, if we leave out of account the normalising factor, by eqn. (13) in § 7, Chap. I,

$$R = \rho^l e^{-\rho/2} \frac{d^{2l+1}}{d\rho^{2l+1}} L_n + h, \quad \rho = 2r \sqrt{\frac{8\pi^2 \mu}{h^2} (-E)}. \quad (1)$$

The Laguerre polynomial L was given by the following equation (14):

$$L_n(\rho) = e^\rho \frac{d^n}{d\rho^n} (\rho^n e^{-\rho}) \quad (2)$$

In the discrete spectrum ρ is real and n an integer. On the other hand, in the continuous spectrum ρ is imaginary, n likewise imaginary and non-integral. For, from (1) it follows for $E = \mu v^2/2 > 0$ that

$$\rho = 2ir \frac{2\pi\mu v}{h} = 2ikr, \quad k = \frac{2\pi}{\lambda}, \quad \lambda = \frac{h}{\mu v} \quad (3)$$

On the other hand, we easily conclude from the connexion between n and Balmer's formula that for $E > 0$

$$n = \frac{Z}{ia k}, \quad a = \frac{h^2}{4\pi^2 \mu e^2} \quad (4)$$

The expression (2) is to be used only for integral n 's. By means of Cauchy's theorem we may, however, easily transcribe it into a formula which holds for any n . For by Cauchy's theorem

$$f(x) = \frac{1}{2\pi i} \int \frac{f(z)}{z-x} dz \quad \text{and hence} \quad \frac{d^n f(x)}{dx^n} = \frac{\Gamma(n+1)}{2\pi i} \int \frac{f(z) dz}{(z-x)^{n+1}}.$$

For arbitrary n 's we therefore have

$$L_n(\rho) = \Gamma(n+1) e^\rho \frac{1}{2\pi i} \int z^n e^{-z} (z-\rho)^{-n-1} dz \quad (5)$$

The integral is here to be taken along a closed path in the complex z -plane, which encloses the point $z = \rho$ and, besides, the point $z = 0$, which, like $z = \rho$ is a branch-point of the integrand in (5) when n is non-integral.

We change the variable of integration to

$$x = z - \rho,$$

and we may write more simply in place of (5)

$$L_n(\rho) = \Gamma(n+1) \frac{1}{2\pi i} \int (x+\rho)^n x^{-n-1} e^{-x} dx \quad (6)$$

the integral now being taken around the branch points $x=0$ and $x=-\rho$. From this we obtain a generally valid integral representation for R in (1), if we replace n by $n+l$ and carry out $(2l+1)$ times the differentiation with respect to ρ . Leaving out a constant factor composed of Γ -functions and adding a power of i that will be convenient in the sequel, we get

$$R = (-i)^l e^{-\rho/2} \frac{1}{2\pi i} \int (x+\rho)^{n-l-1} x^{-n-l-1} e^{-x} dx \quad (7)$$

From the method by which this expression has been derived it is evident (and may also be verified without difficulty) that (7) satisfies the differential equation (3) of p. 60, just like the formula (1) at the beginning.

The integral in (7) is capable of expansion in a convergent series of increasing powers of ρ . We have only to set

$$(x+\rho)^{n-l-1} = x^{n-l-1} \left(1 + \binom{n-l-1}{1} \frac{\rho}{x} + \binom{n-l-1}{2} \frac{\rho^2}{x^2} + \dots \right) \quad (8)$$

This series is absolutely convergent, as the path of integration in the x -plane (Fig. 31) can always be chosen in such a way that $|x| > |\rho|$. We get for the integral in (7) by once again applying Cauchy's theorem:

$$\begin{aligned} & \int \frac{e^{-x}}{x^{n-l+1}} dx + \binom{n-l-1}{1} \rho \int \frac{e^{-x}}{x^{n-l+2}} dx + \binom{n-l-1}{2} \rho^2 \int \frac{e^{-x}}{x^{n-l+3}} dx + \dots \\ &= -2\pi i \left(\frac{1}{(2l+1)!} - \binom{n-l-1}{1} \frac{\rho}{(2l+2)!} + \binom{n-l-1}{2} \frac{\rho^2}{(2l+3)!} \dots \right) \end{aligned}$$

Substituting this in (7) we get the following series, which is convergent for all values of ρ :

$$R = -(-i\rho)^l e^{-\rho/2} \left(\frac{1}{(2l+1)!} - \binom{n-l-1}{1} \frac{\rho}{(2l+2)!} + \dots \right) \quad (9)$$

The closed path of integration of Fig. 31 can be divided into two loops that begin and end at infinity and encircle only one of the two branch points $x=0$ and $x=-\rho$, respectively. Corresponding to this division of the path of integration R splits up into two parts. We set

$$R = \frac{1}{2}(R^1 + R^2) \quad (10)$$

and define R^2 by the loop around $x=0$, R^1 by that around $x=-\rho$.

This resolution is analogous to that known from the theory of Bessel's functions

$$J_n = \frac{1}{2}(H_n^1 + H_n^2).$$

Like H^1 and H^2 so R^1 and R^2 are singular at the origin, while their sums, namely J or R , respectively, behave regularly. This singularity and

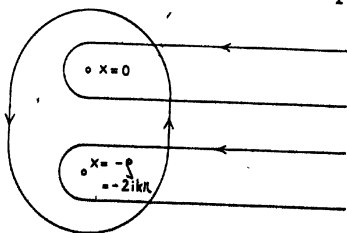


FIG. 31.

regularity, respectively, follow directly from the form of the path of integration: in the case of R^1 and R^2 the path between the two singular points is included in the path of integration which, for $\rho = 0$, passes through these points; in the case of R the path of integration is free.

We investigate the asymptotic behaviour of R^2 . To do this we set, in contradistinction to (8),

$$(x + \rho)^{n-l-1} = \rho^{n-l-1} \left(1 + \binom{n-l-1}{1} \frac{x}{\rho} + \binom{n-l-1}{2} \frac{x^2}{\rho^2} + \dots \right) \quad (11)$$

This expansion is divergent for the more distant parts of the loop-integral, since here obviously $|x| > \rho$. For this reason the following asymptotic development also becomes divergent (semi-convergent).

Substituting in (7) we get in view of (10), if we perform the integration over the loop around $x = 0$,

$$\frac{1}{2}R^2 = \rho^{n-1} e^{-\rho/2} \frac{(-i)^l}{2\pi i} \left(\int \frac{e^{-x}}{x^{n+l+1}} dx + \binom{n-l-1}{1} \frac{1}{\rho} \int \frac{e^{-x}}{x^{n+l}} dx + \dots \right) \quad (12)$$

The integrals here remaining are Γ -functions. For, for arbitrary values of n (even imaginary) the relation

$$\frac{1}{\Gamma(n+1)} = \frac{e^{i\pi n}}{2\pi i} \int \frac{e^{-x} dx}{x^{n+1}} \quad (13)$$

holds. On account of Cauchy's theorem this formula is evidently valid for integral values of n , and for non-integral values it gives the simplest definition of Γ . The path of integration in (13) is the loop of Fig. 31 around $x = 0$ with the condition that, on the upper branch of the path, in $x^n = e^{n \log x}$ the logarithm is to be taken as real. So eqn. (12) becomes

$$\frac{1}{2}R^2 = \frac{\rho^{n-1} e^{-\rho/2} e^{-i\pi(n+\frac{3}{2}l)}}{\Gamma(n+l+1)} \left(1 - \binom{n-l-1}{1} \frac{\Gamma(n+l+1)}{\Gamma(n+l)} \frac{1}{\rho} + \dots \right)$$

and for this we may write

$$\frac{1}{2}R^2 = \rho \frac{e^{-\rho/2 + n \log \rho - i\pi n + \frac{i\pi}{2}l}}{\Gamma(n+l+1)} \left(1 - \binom{n-l-1}{1} \frac{n+l}{\rho} + \dots \right) \quad (14)$$

or on account of the meaning of ρ and n , eqns. (3) and (4),

$$R^2 = C \frac{e^{-\frac{Z}{ak}(kr+a)}}{r} (1 - \dots) \left\{ \begin{array}{l} \alpha = \frac{Z}{ak} \log 2kr \\ C = \frac{1}{ik} \frac{e^{-\frac{\pi}{2} \frac{Z}{ak} + \frac{i\pi}{2}l}}{\Gamma(n+l+1)} \end{array} \right\} \quad (15)$$

This result agrees with the statement in eqn. (4), p. 179; the significance of α compared with that in eqn. (24a), p. 67, has been changed by a factor of $2k$, which, however, since it occurs under the logarithm, may be included in C .

We come to the other component R^1 . In this case it is convenient to rescind the substitution $x = z - \rho$, which led us from (5) to (6) and (7). The loop around $x = -\rho$ then becomes a loop-around $z = 0$. But we must remember that if we have encircled $x = 0$ in the x -plane then, by (7), we have taken up the factor $e^{2\pi i(-n-l-1)} = e^{-2\pi i n}$. We have, therefore, to multiply the integral around the loop $x = -\rho$ by this factor in order

that the sum of both loops, as we demand, may be equivalent to the closed path in Fig. 31. In this way we obtain

$$\frac{1}{2}R^1 = (-i\rho)^l e^{+i/2} \frac{e^{-2\pi i n}}{2\pi i} \int z^{n-l-1} (z-\rho)^{-n-l-1} e^{-z} dz.$$

Since

$$(z-\rho)^{-n-l-1} = (-\rho)^{-n-l-1} \left(1 - \left(\frac{-n-l-1}{1} \right) \frac{z}{\rho} + \dots \right),$$

and taking into consideration the definition of the Γ -function in eqn. (13), we get

$$\frac{1}{2}R^1 = \frac{e^{i/2} e^{-n \log(-\rho) - i\pi n - \frac{i\pi}{2}}}{(-\rho)\Gamma(-n+l+1)} \left(1 - \left(\frac{-n-l-1}{1} \right) \frac{-n+l}{-\rho} + \dots \right) \quad (16)$$

(16) arises out of (14) by exchanging ρ , n and i for $-\rho$, $-n$, $-i$; since ρ and n are purely imaginary, this means that R^1 and R^2 are conjugate imaginaries. Consequently eqn. (15) also holds for R^1 , if we exchange $+i$ for $-i$ in it; this, too, agrees with our earlier assertion in eqn. (4) of p. 179. At the same time it follows from this that the proper function R itself is real.

We can write the expression for R more symmetrically if in place of x we introduce the new variable of integration ξ given by

$$x = \rho(\xi - \frac{1}{2}).$$

We then get from (7)

$$R_l = \frac{(i\rho)^{-l-1}}{2\pi} \int \left(\xi + \frac{1}{2} \right)^{n-l-1} \left(\xi - \frac{1}{2} \right)^{-n-l-1} e^{-\rho\xi} d\xi \quad (17)$$

Here the path of integration (cf. Fig. 32) encircles the two branch points $\xi = \pm \frac{1}{2}$. It is possible but not advisable to contract it on to the real axis from $\xi = -\frac{1}{2}$ to $\xi = +\frac{1}{2}$.

Lastly, we have to deal with the normalisation. Whereas in the discrete spectrum we normalise

$$\int d\tau \psi(E) \psi^*(E) = 1, \quad (18)$$

in the continuous spectrum we must define the normalisation for a finite (even arbitrarily small) interval ΔE by means of the condition

$$\int_{E_1}^{E_2} d\tau \psi(E) \int_{E_1}^{E_2} \psi^*(E') dE' = 1 \quad (19)$$

Here the interval $\Delta E = E_2 - E_1$ is to enclose the point E . In the continuous spectrum the integration (18) would evidently diverge. The integral (19) also diverges, if we reverse the order of integration. To be able to do this we must restrict the volume-integration to the interior of a sphere of very great radius ω and must form

$$\lim_{\omega \rightarrow \infty} \int_{E_1}^{E_2} dE' \int_{r \leq \omega} d\tau \psi(E) \psi^*(E') = 1 \quad (20)$$

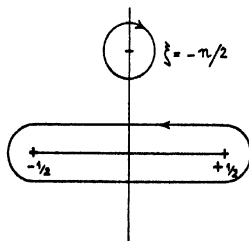


Fig. 32.

From the wave-equations

$$\Delta\psi(E) + \frac{8\pi^2\mu}{h^2}(E - V)\psi(E) = 0,$$

$$\Delta\psi^*(E') + \frac{8\pi^2\mu}{h^2}(E' - V)\psi^*(E') = 0,$$

it follows, by the method of Green's theorem, that

$$\frac{8\pi^2\mu}{h^2}(E - E') \int_{r \leq \infty} \psi(E)\psi^*(E') d\tau = \int_{r \leq \infty} \left(\psi(E) \frac{d\psi^*(E')}{dn} - \psi^*(E') \frac{d\psi(E)}{dn} \right) d\sigma \quad (21)$$

Let us set (N being the normalising factor)

$$\psi = NR_4 P_l^m(\cos \theta) e^{im\phi}, \quad \psi^* = NR_4 P_l^m(\cos \theta) e^{-im\phi} \quad (22)$$

and perform the integration on the right in (21). On account of (23) and (30) in § 6, Chap. I, we get, omitting the suffix l for the present,

$$\frac{4\pi}{2l+1} \frac{(l+m)!}{(l-m)!} N^2 \left(r^2 R(E) \frac{dR(E')}{dr} - r^2 R(E') \frac{dR(E)}{dr} \right)_{r=\infty} \quad (23)$$

Here R can be calculated asymptotically, namely from (15), if we set

$$\arg \Gamma\left(\frac{Z}{ia\hbar} + l + 1\right) = -i\sigma, \quad C_1 = \frac{e^{(-\frac{\pi}{2} \frac{Z}{a\hbar})}}{\left|\Gamma\left(\frac{Z}{ia\hbar} + l + 1\right)\right|},$$

$$R = \frac{1}{2}(R^1 + R^2) = -\frac{C_1}{k\tau} \sin\left(kr + \alpha - \sigma - \frac{\pi}{2}l\right).$$

In agreement with the definition (3) of k we write

$$\frac{8\pi^2\mu}{h^2}E = k^2, \quad \frac{8\pi^2\mu}{h^2}E' = k'^2,$$

and also distinguish the quantities $C, C', \alpha, \alpha', \sigma, \sigma'$ according to the value k or k' . By simple reduction we get for the bracketed expression in (23), except for terms that vanish with r , when $r = \infty$,

$$\frac{C_1 C_1'}{2kk'} \{ - (k' + k) \sin p + (k' - k) \sin q \},$$

where

$$p = (k' - k)\omega + (\alpha' - \alpha) - (\sigma' - \sigma),$$

$$q = (k' + k)\omega + (\alpha' + \alpha) - (\sigma' + \sigma).$$

Thus from (21) it follows, if we divide out by the factor on the left-hand side, that is, by $k^2 - k'^2$, and take (22) into account,

$$\int_{r \leq \infty} \psi(E)\psi^*(E') d\tau = \frac{4\pi}{2l+1} \frac{(l+m)!}{(l-m)!} N^2 \frac{C_1 C_1'}{2kk'} \left\{ \frac{\sin p}{k' - k} - \frac{\sin q}{k' + k} \right\}.$$

By (20) this expression is to be multiplied by $dE' = \frac{h^2}{4\pi^2\mu} k' dk'$ and to be

integrated from E_1 to E_2 . Let the corresponding limits for k' be k_1 and k_2 . Thus by (20) we have

$$1 = \frac{4\pi}{2l+1} \frac{(l+m)!}{(l-m)!} N^2 \frac{h^2}{8\pi^2\mu} \frac{C_1}{k} \int_{k_1}^{k_2} C_1' dk' \left\{ \frac{\sin p}{k' - k} + \dots \right\} \quad (24)$$

The unwritten second term clearly contributes an amount which is vanishingly small for a vanishingly small interval of integration, as it is finite throughout this interval. The integral over the first term of the bracket can be simplified to

$$C_1 \int_{p_1}^{p_2} \frac{dp}{p} \sin p \quad . \quad . \quad . \quad (25)$$

where p_1 and p_2 are the values of p that correspond to the values $k' = k_1$ and $k' = k_2$. Since k lies within the interval (k_1, k_2) , we find $p_1 = -\infty$, $p_2 = +\infty$ for $\omega = \infty$. Thus (25) is equal to πC_1 and (24) gives

$$N^2 = (2l+1) \frac{(l-m)!}{(l+m)!} Q, \quad Q = \frac{2\mu}{\hbar^2} \frac{k}{C_1^2} = \frac{2\mu k}{\hbar^2} \frac{\pi Z}{e a k} \left| \Gamma\left(\frac{Z}{ia k} + l + 1\right) \right|^2 \quad (26)$$

The expression (17) for the radial component R of the proper function and the expression (15) for the partial functions R^1, R^2 are to be multiplied by the normalising factor N so determined, in accordance with (22). Our calculation of N is practically identical with the method first developed by E. Fues.*

2. Concerning the One-sidedness of Photo-electric Emission. Calculation of the Integrals K_1 and K_2

The integrals K_1 and K_2 are defined on p. 186. According to (17)† we have, if we add the normalising factor given by (26),

$$(\rho)^2 R_1 = \frac{N_1}{2\pi} \int (\xi + \frac{1}{2})^{n-2} (\xi - \frac{1}{2})^{-n-2} e^{-\rho \xi} d\xi,$$

$$(\rho)^2 R_2 = \frac{N_2}{2\pi} \int (\xi + \frac{1}{2})^{n-2} (\xi - \frac{1}{2})^{-n-2} e^{-\rho \xi} d\xi,$$

$$n = \frac{Z}{ia k}, \quad \rho = 2ikr,$$

$$\frac{N_2}{N_1} = \sqrt{\frac{5 \cdot 2!}{3 \cdot 3!}} \left| \frac{\Gamma(n+3)}{\Gamma(n+2)} \right| = \sqrt{\frac{5}{9}} |n+2| \quad . \quad . \quad (27)$$

In K_1 and K_2 we exchange the integrations with respect to r and ξ and perform the integration with respect to r . In each case we get

$$\int_0^\infty e^{-\frac{Zr}{a} - 2ikr\xi} dr = \frac{1}{\frac{Z}{a} + 2ik\xi} = \frac{1}{2ik\left(\xi + \frac{n}{2}\right)} \quad . \quad . \quad (28)$$

And so

$$\left. \begin{aligned} (2k)^2 \frac{K_1}{N_1} &= \frac{1}{2\pi i} \int \frac{(\xi + \frac{1}{2})^{n-2} (\xi - \frac{1}{2})^{-n-2} d\xi}{\xi + n/2} \\ - (2k)^2 \frac{K_2}{N_2} &= \frac{1}{2\pi i} \int \frac{(\xi + \frac{1}{2})^{n-2} (\xi - \frac{1}{2})^{-n-2} d\xi}{\xi + n/2} \end{aligned} \right\} \quad . \quad (29)$$

As in Fig. 32 the integrations are to be performed around the two branch-points $\xi = \pm \frac{1}{2}$ with the exclusion of the newly-added singular point

* Ann. d. Phys., 81, 281 (1926), particularly § 8.

† The references are to equations in the Notes unless the contrary is stated.

$\xi = -n/2$. The latter follows from (28), since here the integration has a meaning only if

$$\frac{Z}{a} > 2k\xi_2, \text{ that is, } 2\xi_2 < |n|,$$

holds for the whole path of integration, where ξ_2 denotes the imaginary part of ξ . The integrations in (29) may be performed in Cauchy's manner. The integrals vanish at ∞ like ξ^{-5} and ξ^{-7} respectively; hence with continued extension of the path of integration only the residues at the point $\xi = -n/2$ remain.

In this way we get

$$\begin{aligned} (2k) \frac{K_1}{N_1} &= \left(-\frac{n}{2} + \frac{1}{2}\right)^{n-2} \left(-\frac{n}{2} - \frac{1}{2}\right)^{-n-2}, \\ (2k) \frac{K_2}{N_2} &= \left(-\frac{n}{2} + \frac{1}{2}\right)^{n-2} \left(-\frac{n}{2} - \frac{1}{2}\right)^{-n-2}, \\ 2k \frac{K_2}{K_1} &= \frac{4}{1-n^2} \frac{N_2}{N_1}. \end{aligned}$$

From (15), adding the normalising factors, we then form

$$\frac{R_2^2}{R_1^2} = i \frac{\Gamma(n+2)}{\Gamma(n+3)} \frac{N_2}{N_1} = \frac{i}{n+2} \frac{N_2}{N_1}.$$

Hence, on account of (27)

$$\frac{K_2 R_2^2}{K_1 R_1^2} = \frac{2i}{k} \cdot \frac{5}{9} \frac{|n+2|^2}{(1-n^2)(n+2)} = \frac{2i}{k} \cdot \frac{5}{9} \frac{2-n}{1-n^2}. \quad (30)$$

This expression is to be substituted in eqn. (18) of p. 186, and with $n = \frac{Z}{iak}$ it immediately yields the value of Ψ there given.

3. The Integration over the Continuous Spectrum Transition from the Proper Function ψ to the Partial Functions ψ^1, ψ^2

We must here deal with a somewhat awkward point which has not been sufficiently taken into consideration in the German original of this book. For a more complete treatment of the subject in question the reader is referred to a paper which is shortly to appear in the *Annalen der Physik*. We must here restrict ourselves to indicating the line of reasoning.

The integration with respect to E' in eqn. (1) of p. 178 is to be performed between $E' = \epsilon_0 = \mu c^2$ to $E' = \infty$, but it must not be taken along the real axis, because there the denominator vanishes for

$$E' = E_k + h\nu. \quad (31)$$

Hence the path of integration must avoid this point by escaping into the complex domain. In place of E' the wave-number k' suggests itself as a convenient complex variable, which is associated with the kinetic energy $\epsilon' = E' - E_0$

$$k'^2 = \frac{8\pi^2\mu}{h^2} (E' - E_0) = \frac{8\pi^2\mu}{h^2} \epsilon'. \quad (32)$$

We then have for the differential in I ,

$$\frac{dE'}{E_k - E' + h\nu} = \frac{2k'dk'}{k^2 - k'^2}.$$

The wave-number k which here occurs in the denominator is associated with the Einstein energy ϵ (eqn. (5) of p. 180) just as k' is associated with the kinetic energy ϵ' . In the complex k' -plane the path of integration leads, on account of (32), from $k' = 0$ to $k' = \infty$, the singular point k being avoided, cf. Fig. 33.

From the mathematical standpoint we may make the path of integration deviate upwards as well as downwards. This arbitrariness corresponds to an ambiguity (*Vieldeutigkeit*) which is always present in a continuous spectrum and can be avoided only by means of a special condition* at infinity "the condition of radiation." According to this condition, if we write the time-factor as in I, p. 178, with a positive exponent, then we must choose our route of escape in the positive imaginary half-plane. (We could just as well have written the exponent with a negative i and then we should have had to choose the dotted path in Fig. 33.)

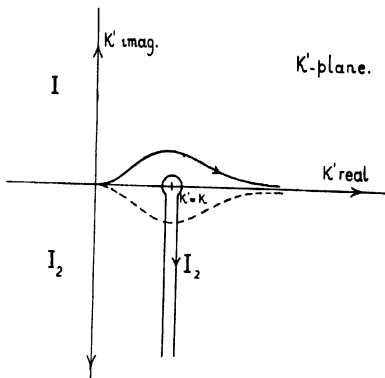


FIG. 33.

We now split ψ up into its two partial functions, in a manner analogous to that in eqn. (10) of Note I:

$$\psi = \frac{1}{2}(\psi^1 + \psi^2),$$

and we get for the integral I ,

$$I = \int (\psi^1 + \psi^2) \frac{Ak'dk'}{k^2 - k'^2} \cdot e^{\frac{2\pi i}{h}(E_k + h\nu)t}.$$

On account of the asymptotic method of representing the partial functions, eqns. (15) and (16) of Note I,

$$\psi^1 \sim e^{ikr + \dots}, \quad \psi^2 \sim e^{-ikr + \dots}. \quad (33)$$

ψ^1 vanishes at infinity in the positive imaginary half-plane, ψ^2 at infinity in the negative imaginary half-plane. It is true that the factor A , in virtue of its being defined by a volume integral contains the factor ψ implicitly. But the value of r in this space-integral—we call it r' and have $0 < r' < \infty$ —is in general smaller than the value of r in ψ^1 , ψ^2 which occur explicitly; and as we shall allow r to go asymptotically to infinity, only the quantities ψ^1 , ψ^2 will be decisive for the behaviour at

* Cf. A. Sommerfeld, Jahresbericht d. deutsch. Math. Vereinigung, Vol. 21, p. 309, Part B, Chap. II, or the above-mentioned paper in the Ann. d. Phys.

infinity in the k' -plane. We divide I into two parts I_1 and I_2 and lead round the part I_1 formed with ψ^1 in the positive imaginary half-plane, the part I_2 formed with ψ^2 in the negative imaginary half-plane. The latter is arrested at the singular point $k' = k$. In addition, we have the two integrations of I_1 and I_2 with respect to the positive and the negative imaginary axis. But these integrations vanish exponentially, according to (33), if r is sufficiently great. Then there remains only the residue at the point $k' = k$, so that we can write

$$I = \pi i \psi^2(E) A(E) \text{ or, better, } I = \pi i \sum \sum \psi^2(E) A(E). \quad (34)$$

E is the energy-value defined by (31) and belonging to $k' = k$. The double summation ranges, as in eqns. (14) and (18) on pp. 182 and 186, over all values of the quantum numbers l, m so long as they are not excluded owing to the factor A vanishing in any way.

In conclusion we have some remarks to make about integral II in eqn. (1), p. 178. The denominator in this integral does not vanish between $E_0 \leq E' < \infty$, since

$$E' = E_k - h\nu = E_0 - J - h\nu < E_0$$

holds for the zero-point of this denominator, where J denotes the work of ionisation of the state k . If we also here introduce the wave-number k' belonging to the kinetic energy $\epsilon' = E' - E_0$, the choice of the path of integration in the k' -plane, Fig. 33, becomes immaterial, since there is no singular point to the right of the imaginary axis. Thus if we divide II up into Π_1 and Π_2 , these partial integrals consist solely of the integrals with respect to the positive and negative imaginary axis. They vanish just as, before, the partial integrals I_1 and I_2 taken along the imaginary half-axes vanished. The fact that of the two integrals I and II only the residue calculated in (34) remains is clearly to be interpreted as a resonance phenomenon. We were therefore justified in not taking the component part II into account in our discussion of the photo-electric effect.

INDEX

German Expressions and their English Equivalents

(The Numbers Refer to the Pages on which the Expressions Occur)

- Adjungiert*, adjoint, 104.
Anregungsstärken, excitation intensities, 47.
Ansatz, assumption, expansion or substitution, 8.
Anschaulich, picturesque (intuitive), 47.
Anschaulichkeit der Vorstellung, vividness of the mental pictures, 76.
Aufenthaltsmöglichkeit, possibility of sojourn, 256.
Aufspaltungen, separations, 115.
Ausserwesentlich singuläre Stelle, pole, 9.
Austausch-Entartung, exchange degeneracy, 18, 227, 239.
Bahnimpuls, orbital moment of momentum, 96.
Eigenwert, proper or characteristic value, 7, 42.
Eigenzustand, proper state, 42.
Eindeutig, uniform (single-valued), 7, 8, 101.
Eindeutigkeit, uniformity (mathematical), 41.
Eindeutigkeitsprinzip, principle of uniformity, 122.
Einkristall, single crystal, 205.
Elektronenstoss, collisions caused by electrons, 47.
Ersatzelektron, substitute electron, 220.
Flecke, spots (points), 205.
Halbierungskegel, bisecting cone, 187.
Halbierungswinkel, bisecting angle, 192.
Hilfsgrösse, auxiliary quantity, 84.
Hohlraum, enclosure, cavity, 28, 77.
Interkombinationsverbot, law forbidding combinations between certain states, 280.
Knotenlinie, line of nodes, 127, 150.
Knotensahl, number of nodes, 187.
Kreiselmoleküle, molecular tops, 128.
Ladungswolke, charge cloud, 75.
Leuchtelektron, radiating electron, 95.
Magnetfeld, magnetic field, 97.
Mulde (energy-) pocket, 208.
Norm, square of the absolute value of matrix elements, 84, 58, 84.
Ortsfunktion, space function, 172.
Oszillatorenstärke, number or quantity of oscillators, 170.
Partiell verkehrt, partially inverted, 284.
Pauli Verbot, Pauli's exclusion principle, 85.
Pendelbahnen, pendulum orbits, 62.
Platzwechsel or *Austausch*, exchange of position, 229.
Resonanz-Entartung, resonance degeneracy, 227.
Richtungsquantelung, quantising of direction, 97.
Ruhenergie, rest energy, 41.
Selbstadjungiert, self-adjoint, 105.
Spielraum, range of play, 256.
Spitzenzähler, point-counter, 191.
Stossquerschnitt, effective cross-section in collisions, 97.
Überlichtgeschwindigkeit, velocity exceeding that of light, 121.
Umdrehungsphase, phase of rotation, 202.
Umkehrung, inverse, 181.
Ungenauigkeitsrelation, uncertainty relation, 84, 122.
Vertauschungsrelation, commutation law, 82.
Vieldeutigkeit, many-valued character, 297.
Viererpotential, four potential, 100.
Vollständigkeitsrelation, completeness relation, 147, 221.
Voreilwinkel, forward angle, 191.
Wahrscheinlichkeitspaket, probability packet, 256.
Wechselsatz, Alternation Law, 94.
Zustände, states, 120.
Zustandsgrössen, phase or state parameters, 80.
Zustandskonstante, state-constant, 42.
Zustandsverteilung, distribution of state, 167.
Zweigliedrig, having two terms, 28.

INDEX

A

Aage, 128.
Action function, 3.
Addition theorem of spherical harmonics, 86.
Adiabatic principle, 62.
Adjoint, 104.
Alkali character of Röntgen terms, 94.
— like doublet character of hydrogen lines, 94.
Alternation law, 94.
Andrade, 212.
Angular co-ordinates, 69, 270.
Anomalous Zeeman effect, 2.
Anti-parallel transitions, 57.
Areas, law of, 251.
Associated spherical functions, 12.
— — harmonics, 12.
Auger, 191, 192.
Azimuthal quantum number, 62, 78, 95, 254, 287.

B

Balmer series, 76, 79, 81, 83.
— terms, 68, 69, 114.
Band spectra, 1, 24, 99.
— — Kratzer's theory, 21.
Barlow, 191.
Barnett, 220.
Bearden, 220.
Bechert, 11.
Beck, 176.
Bessel functions, 18, 150.
Bethe, 205, 211.
Bige, 47.
Bleeker, 88.
Bochner, 270.
Bohr, 18, 68, 86, 255.
— circles, 70.
— magneton, 261, 265.
Bohr's frequency condition, 29.
Bonhoeffer, 240.
Born, 80, 83, 99, 193, 200.
Bothe, 190, 217.
Boundary conditions, 6, 62, 115.
Bowen, 118.
Brackett series, 83.
Bragg, 209.
Breit, 213.
Brillouin, 134.
Broglie, *de*, 7, 41, 122, 205, 208.
— — wave, 7, 83, 89.
— — wave-length, 7, 244.

Bubb, 191.
Burger, 46, 59.

C

Calculus of matrices, 30, 81.
Canal rays, 97.
Cauchy, 36.
Causality, 188.
Characteristic (see proper and *eigen*).
— equation, 9.
— state, 42.
— values, 42.
Charge-cloud, 69, 75, 77, 83, 98.
— density, 74.
Chladni's figures, 149.
Circular membrane, 150.
Cloud chamber, Wilson's, 178, 191, 200.
C-numbers, 243.
Collisions, 47, 96, 192, 201.
Combination principle, 29.
Commutative law, 31, 32.
Completeness relation, 147, 221.
Compton effect, 175, 191, 212, 213.
Condon, 121.
Conjugate imaginaries, 42.
Conservation of charge, 89.
— — energy and momentum, 213 *et seq.*
Continuity, equation of, 107.
Continuous spectrum, 59, 66, 178.
Core of atom, 163.
Correspondence principle, 30, 59.
Courant and Hilbert, 8, 147.
Crystal lattices, 98.
Crystals, 268.
— two-dimensional cubic, 268.
Curl, four-dimensional, 260.
Current, 171.
— density, 90.
— matrix, 90.
Cyclic co-ordinates, 8.
Czerny, 17.

D

Darwin, 288, 286.
Davison, 205 *et seq.*
Debye, 4, 205, 212.
Debye-Scherrer rings, 212.
Deflection of α -rays, 193, 198.
Deformation of atom, 163.
Degeneracy, 18, 159, 227.

Degenerate proper value problems, 18, 68, 148.

Dennison, 124.

Density, electric, 167.

— of charge, 44, 45.

— — mass, 44.

Deslandres term, 204.

Diagonal matrix element, 222.

Diamagnetic, 93.

Diatomic molecules, 27.

Diffraction of electrons, 7.

— problems, 150, 192, 205.

Diffuse series, 81.

Diffusion coefficients, 43.

— equation, 43.

Dipole moments, 88.

Dirac, 47, 83, 86, 118, 213, 219, 231, 234, 243, 255 *et seq.*

Discrete spectrum, 59, 62.

Discriminant, 129.

Dispersion formula, 169.

Div, 89, 90.

Donder de, 118.

Dorgelo, 46, 59.

Doublets, screening and relativity, 116, 288.

Dungen, van den, 118.

E

Eckart, 51.

Eddington, 190.

Effective cross-section, 97.

Ehrenfest, 244.

Eikonal, 4.

Einstein, 102.

— photo-electric equation, 88, 178, 180.

— probabilities of transition, 170.

Electron in electromagnetic field, 259.

— magnetism, 115.

— rays, 198.

— under no force, 257.

Electronic emission, 184.

— spin, 115, 116, 231.

Elliptic orbits, 60, 178.

Elsasser, 193, 200.

Energetics, 83.

Energy, 241.

— equation of, 8, 241.

— levels, 141.

— matrix, 86.

— momentum theorem, 110.

— of position, 42.

— parameters, 86.

Epstein, 121, 154.

Equatorial quantum number, 78.

Essential singularities, 9, 67.

Eucken, 240.

Eulerian angles, 134, 127.

Ewald, 211.

Exchange degeneracy, 18, 227, 239.

— of position, 229.

Excitation intensities, 47.

Exclusion principle, Pauli's, 210, 233.

F

Fermi, 92, 201.

Fine-structure, 114, 116, 233.

— constant, 100, 284.

— of helium, 231.

Fock, 118.

Formaldehyde, 124.

Försterling, 117.

Forsyth, 11.

Fourier series, 80, 147.

Four-current, 104, 107.

— -dimensional curl, 260.

— -group, 270.

— -potential, 100.

— -vector, 92, 260.

Fowler, 59.

Frazer, 97, 98.

Frequency condition, Bohr's, 1, 38, 45.

Fresnel, 209.

Fues, 21, 27, 52.

Functional determinant, 129.

G

Gauss's error curve, 16.

— theorem, 109.

Geiger, 191.

General harmonics, 9.

Generalised co-ordinates, 6.

Gerlach and Stern, 37, 96, 98.

Germer, 205 *et seq.*

Geometrical-optics, 2.

g-formula, Landé's, 2, 9, 62, 255.

Gordon, 110, 118, 200, 213.

Goudsmit and Uhlenbeck, 85, 117, 232, 261, 265.

Gradient, 89, 90.

Green's functions, 106.

— theorem, 43, 83, 90, 104, 106, 108, 144, 195.

Ground states, 96, 170, 189.

Group, 89.

— mean amplitude of, 89.

— quantum number, 93, 94.

— velocity, 121.

H

Half-integers, 1, 2, 16, 17, 23, 24, 86, 94.

Hamiltonian functions, 3, 32, 86, 111.

— mechanics, 2.

Hamilton's differential equation, 8, 257.

— variation principle, 110.

Hankel solutions, 14.

Hansen, 117, 226.

Harmonic oscillator, 137.

Harmonics, 9.

Hartree, 240.

Heine, 56.

Heisenberg, 29, 30, 37, 83, 47, 83, 169, 225, 233, 238, 239, 255, 256, 257.

Heitler, 240.

Helium, ortho- and para-, 226.

— problem, 2, 144, 225.

Helium, spectrum, 226.
 Henri, Victor, 28.
 Hermitean matrices, 16, 31, 43, 49.
 — polynomials, 46.
 Hilbert, 36.
 Hölml, 59.
 Houston, 204, 226.
 Hull, 204.
 Huygen's principle, 198.
 Hydrodynamic theorem of continuity, 89.
 Hydrogen spectrum, 290.
 Hylleraas, 112.
 Hyperbolic orbits, 65, 178.

I

Indefiniteness, 255.
 — relation of, 255.
 Inner electrons, 229.
 — quantum number, 94.
 Intensities, 48, 79.
 Intensity, 29, 168.
 — of oscillators, 170.
 Interference of electron waves, 205.

J

Jacobian polynomials, 133.
 Jordan, 30, 63, 84, 255.

K

Kallmann, 220.
 Kalusa, 102.
 Kellner, 112.
 Kepler, equations of, 282.
 — problem, 59, 68, 73, 92, 139.
 — third law, 346.
 Kikuchi, 212.
 Kirchner, 191, 200, 223.
 Klein, 102, 118, 149, 164, 218.
 Kokubu, 163.
 Kramers, 169, 171.
 Kratzer, 21, 24.
 Krishnan, 174.
 Kronig, 124.
 Kudar, 118.
 Kunsman, 205.
 Kupper, 52.

L

Ladenburg, 170.
 Lagrange co-ordinates, 6.
 — function of wave-mechanics, 110.
 Laguerre polynomials, 25, 50, 63, 69, 70, 153.
 Landé's g -formula, 2, 62, 94, 99, 255.
 Landsberg, 174.
 Larmor's precession, 108, 131.
 — theorem, 130.
 Laue, von, 104, 211.

Law of uncertainty, 88.
 Legendre functions, 9.
 Legendre's polynomials, 11, 73.
 Lewis, 88.
 Light corpuscles, 8.
 Linear, grating, 202.
 — harmonic oscillator, 14, 48.
 — independence, 148.
 Lines of nodes, 151.
 London, 59, 88.
 Lorentz, 119, 120.
 — transformation, 100.
 — triplet, 123.
 Lyman series, 79, 81.

M

Mach, 38.
 Macro-mechanics and micro-mechanics, 2, 5, 77.
 Macroscopic and microscopic, 1.
 Magnetic change of wave-number, 122.
 — moments, 88.
 — quantum number, 78, 86, 98, 232.
 Magnetron, Bohr, 86, 92, 96.
 Mandelstam, 174.
 Manneback, 124.
 Mark, 163, 219.
 Matrices with two rows, 267.
 — — four rows, 276, 277.
 Matrix, 30.
 — composite, 269.
 — elements, 45, 77.
 Maxwell, 9, 84.
 Mecke, 99.
 Medium transitions, 57.
 Membrane, proper vibrations of, 146.
 — quadratic, 149, 227.
 Microscope, γ -ray, 256.
 Millikan, 118.
 Möglich, 259.
 Moment, 45.
 — electric, 168.
 — magnetic, 259.
 — mechanical, of electron, 262.
 — of first and higher orders, 52.
 Momentum, 240, 244, 251.
 — conservation of, 244.
 — matrix, 176.
 — moment of, orbital, 96.
 — operator, 243.
 — specific, 242.
 Moseley, 116.
 Mulliken, 24.
 Multipole moments, 88, 98.
 Negative dispersion terms, 171.
 Neumann, von, 240, 259.
 Neutral transitions, 57.
 Nickel, 205, 206.

N

π -orbits, 75.

Nodes, 6, 70, 187.

Non-coherent scattering, 173.

Normalisation, 9, 287.

— general condition of, 44, 48.

— persistence of, 89.

Nucleus, motion of, 75.

Nuttall, 191.

O

Observable quantities, 37.

Oppenheimer, 199.

Orbital moment of momentum, 96.

Orbits, electronic, 38, 62, 88, 235.

Ordinary points, 9.

Orthogonality, condition of, 44, 49, 64.

Oscillating rotator, 26.

Oscillator, 37.

— and rotator, 14.

P

Parallel transitions, 57.

Paramagnetic, 93, 96.

Paschen-Back effect, 2, 117.

Paschen series, 82, 83.

Pauli, 59, 62, 81, 82, 83, 233, 270.

— exclusion principle, 85, 86, 93, 210, 233.

Pauling, 74.

Pendulum orbits, 62.

Perron, 73.

Persistence of normalisation, 89.

Perturbations, theory of, 144.

Phase, 39.

— functions, 52.

— of rotation, 202.

— parameters, 30.

— quantities, 29.

— velocity, 39, 208.

Photo-electric, effect, 26, 176, 295.

— — emission, one-sidedness of, 295.

— — equation, 33.

Photo-electrons, 180.

Photon, 256.

Pickering series, 76.

Planck's constant, 5.

Pockels, 149.

Point-counter, 190.

Point-mass under no forces, 38.

Polar crystals, 98.

Polarisation of atoms, 164.

— rule of, 55.

Poles, 9.

Polynomial method, 11, 21, 64, 124, 139, 162.

Polynomials, 73.

— Hermitean, 16.

Principal quantum number, 61, 78, 156, 236.

— series, 81.

Principle of variation, 100.

Pringsheim, 174.

Probability: amplitudes, 84, 256.

Probability of excitation, 170.

— packet, 256.

Problem of three bodies, 239.

Proper functions, 7, 42, 59, 65, 236, 290, 296.

— value parameters, 7, 9, 42, 59, 65.

Protons, 97.

Q

Quadratic membrane, 149.

Quantising of direction, 97.

Quantum condition, 1.

— numbers, 1, 61, 62, 78, 93, 94, 233.

— — meaning of, 69.

— — four, 85.

— oscillation, 23.

— rotation, 23.

— weight, 53, 96.

Quaternion unit, 270.

q -matrix, 46.

q -numbers, 243.

R

Rademacher, 124.

Radial condition of orthogonality, 64.

— quantum number, 61, 78.

Radiating electron, 95.

Radiation, free, 163.

— pressure, 190.

Radiationless states, 37, 45.

Raleigh, Lord, 145.

Raman effect, 174, 175.

Recurrence formulae, 11, 16, 54, 118, 182.

Refractive index, 104.

Regular points, 9.

Reiche, 124.

Relativistic generalisations, 100.

— Kepler-problem, 62, 112.

Repulsive forces in crystal lattices, 98.

Resonance catastrophe, 146.

— degeneracy, 227.

Rest energy, 41.

Resultant mass, 28, 76.

Ricciate equation, 135.

Richardson effect, 208, 210.

Riemann, 37.

Ritz, 112.

Röntgen rays, photo-electric emission of, 187.

— terms, 117.

Rosen, 174.

Rotator, 14.

— in space, 16, 52.

— in a plane, 13.

— oscillating, 20.

Ruark, 244.

Runge, 4.

Rupp, 212.

Rutherford, 193, 197, 212.

— formula for scattering, 199.

Rydberg constant, 61, 68, 76.

S

- Scherrer, 205.
 Schrödinger, 43, 47, 51, 59, 66, 81, 110, 144,
 152, 162, 174, 218, 246, 288.
 — papers, 2, 26, 27, 77, 81, 117, 144, 172.
 Schütz, 98.
 Schwarzschild, 18, 154.
 Screening, 235.
 Sectorial areas, law of, 246.
 — harmonics, 12.
 Selection rules, 46, 49, 78, 79.
 Self-adjoint, 105, 127, 144.
 Separation constant, 9.
 — of variables, 19.
 Separations, 115.
 Series of spectral lines, 81.
 Sharp series, 82.
 Shells, electronic, 75, 84.
 Siegbahn, 117.
 Simon, 218.
 Singular-points, 9.
 Slater, 112, 118.
 Slowly varying quantities, 4, 13.
 Smekal transitions, 167, 213.
 Sommerfeld, 94, 106, 118, 205, 210, 290.
 — and Runge, 4.
 Spatial polar co-ordinates, 59.
 — quantising, 18.
 Specific current, 91.
 Spherical harmonics, 8, 50, 58, 86.
 — — associated, 52.
 — waves, 180.
 Spin, electron, 85, 118, 232.
 Stark effect, 153.
s-state, 115, 151.
 State constant, 42.
 Static octet, 88.
 Stationary currents, 88, 91.
 — states, 1, 29.
 Statistical view of charge cloud, 83, 188,
 199.
s-terms, 85, 94.
 Stern, 97.
 Stratton, 222.
 Stripped electrons, 118.
 Strong transitions, 57.
 Sturm-Liouville-problems, 8, 160.
 Superfluous degrees of freedom, 18.
 Surface harmonics, 9.
 — integrals, 242.
 Symmetrical tensor, 246.
 — tops, 123.

T

- Takamine, 168.
 Tensor, 246.

- Terms, 258.
 Tesseral harmonics, 9.
 Tetrode, 259.
 Thomson, G. P., 210.
 Time-equation, 5, 88, 43, 88.
 Transitions between states, 87, 52, 90, 178.
 — parallel, anti-parallel and neutral, 57.
 — spontaneous, 47.
 — strong, weak, and medium, 57.
 Trumphy, 82.

U

- Uncertainty, relation, 122, 255.
 — law of, 88, 84.
 Undisplaced line of Compton effect, 228.
 Uniformity, 41.
 Uniqueness, principle of, 122.
 Unit matrix, 44.
 Unsold, 85, 86, 87, 99, 118, 226.

V

- Values, continuous and discontinuous of
 spectrum, 13.
 Variation, principle of, 108.
 Vibration equation, 43.
 Virial, law of the, 249.

W

- Waller, 162.
 Wave-equation, 5.
 — function, 5.
 — group, 41.
 — mechanics, 5.
 — number, 8.
 — packet, 41, 255, 256.
 — theory, 2.
 Weak transitions, 57.
 Weight function, 73, 79.
 — of a state, 58, 96.
 Well-defined states, 85.
 Wentzel, 134, 162, 190, 198, 218, 215, 221.
 Weyl, 86.
 Wierl, 168.
 Wigner, 240.
 Williams, 190.
 Wilson, 177.
 Woo, 219.
 Wood, 174.
 Work of dissociation, 24.
 Wrede, 97.

Z

- Zeeman effect, 2, 55, 62, 92, 94, 119, 128.
 Zimmermann, 164.
 Zonal harmonics, 11.

